

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 03:11:06 ; Search time 918.465 Seconds
(without alignments)
9617.545 Million cell updates/sec

Title: US-10-609-078-7

Perfect score: 1455

Sequence: 1 gaattcgacagcgccac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1400	96.2	1755	US-10-425-115-18040	Sequence 18040, A
2	1394.2	95.8	1444	US-10-425-114-15440	Sequence 15440, A
3	1394.2	95.8	1450	US-10-425-114-25244	Sequence 25244, A
4	1196.8	82.3	1706	US-10-425-115-18039	Sequence 18039, A
5	1179	81.0	1649	US-10-425-114-24871	Sequence 24871, A
6	1113.4	76.5	1537	US-10-767-701-13951	Sequence 13951, A
7	1047	72.0	1361	US-10-425-114-25845	Sequence 25845, A
8	1047	72.0	1490	US-10-425-114-22598	Sequence 22598, A
9	1046	71.9	1519	US-10-425-115-18041	Sequence 18041, A
10	753	51.8	2471	US-10-739-930-4858	Sequence 4858, Ap
11	704.8	48.4	1569	US-10-437-963-37229	Sequence 37229, A

12	571	39.2	571	16	US-10-074-473-2	Sequence 2, Appli
13	479	32.9	1691	17	US-10-424-599-55699	Sequence 55699, A
14	476.2	32.7	1598	17	US-10-424-599-81927	Sequence 81927, A
15	416.6	28.6	483	17	US-10-424-599-118531	Sequence 118531, A
16	388	26.7	388	16	US-10-074-473-3	Sequence 3, Appli
17	332.8	22.9	1674	17	US-10-424-599-98272	Sequence 98272, A
18	286	19.7	287	9	US-09-294-0938-571	Sequence 571, App
19	245.2	16.9	1029	17	US-10-424-599-81933	Sequence 81933, A
20	227	15.6	316	18	US-10-425-115-87023	Sequence 87023, A
21	223.4	15.4	384	18	US-10-425-115-81982	Sequence 41982, A
22	221.8	15.2	585	18	US-10-021-323-2492	Sequence 2492, Ap
23	217.8	15.0	534	18	US-10-021-323-14417	Sequence 14417, A
24	216.4	14.9	272	9	US-09-294-0938-4704	Sequence 4704, Ap
25	201.6	13.9	291	9	US-09-294-0938-3135	Sequence 3135, Ap
26	201.6	13.9	851	17	US-10-424-599-75856	Sequence 75856, A
27	194.4	13.4	528	18	US-10-021-323-14351	Sequence 14351, A
28	186.2	12.8	329	18	US-10-425-115-137486	Sequence 137486, A
29	183.8	12.6	573	18	US-10-021-323-2546	Sequence 2546, Ap
30	182.8	12.6	275	18	US-10-425-115-60448	Sequence 60448, A
31	180.6	12.4	556	17	US-10-424-599-98270	Sequence 98270, A
32	175.6	12.1	523	18	US-10-021-323-15961	Sequence 15961, A
33	163	11.2	1555	18	US-10-335-053-99	Sequence 99, Appl
34	163	11.2	1776	18	US-10-357-930-22330	Sequence 22330, A
35	163	11.2	1776	18	US-10-357-930-28174	Sequence 28174, A
36	156.4	10.7	1264	17	US-10-264-049-918	Sequence 918, Appl
37	156.4	10.7	1366	16	US-10-287-218-33	Sequence 33, Appl
38	156.4	10.7	1366	18	US-10-474-291-33	Sequence 33, Appl
39	156.4	10.7	2053	10	US-09-983-802-45	Sequence 45, Appl
40	156.4	10.7	2053	10	US-09-984-490-45	Sequence 45, Appl
41	156.4	10.7	2053	11	US-09-973-278-82	Sequence 82, Appl
42	142.2	9.8	409	9	US-09-770-423-621	Sequence 621, App
43	129.2	8.9	481	10	US-09-770-961-329	Sequence 929, App
44	108.2	7.4	432	17	US-10-424-599-49636	Sequence 49636, A
45	106.8	7.3	479	17	US-10-424-599-25942	Sequence 25942, A

ALIGNMENTS

RESULT 1

US-10-425-115-18040

; Sequence 18040, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 18040

; LENGTH: 1755

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_116455C.1

US-10-425-115-18040

Query Match 96.2%; Score 1400; DB 18; Length 1755;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 1412; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 24 AGTCATCGTCTGTCGTCGCCCGCAGAGCGCGAGCGCCCAAGCGCGAGGGCGCCCAT 83

DB 194 AGTCATCGTCTGTCGTCGCCCGCAGAGCGCGAGCGCCCAAGCGCGAGGGCGCCCAT 253

QY 84 GSCCGCGTCGATCTGGAGCAAGCCCAAGCGCGCTTCGTCGACGACGCTTCGAGCT 143

DB 254 GSCCGCGTCGATCTGGAGCAAGCCCAAGCGCGCTTCGTCGACGACGCTTCGAGCT 313


```
QY 444 ATCTAGCCAGCCACAGTAAGATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
Db 445 ATCTAGCCAGCCACAGTAAGATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 504
QY 504 GGAGGATTTTCAAAATATGAGAAATACACACACAGTGTAGAGACCCCAAGCAAAACCAA 563
Db 505 GGAGGATTTTCAAAATATGAGAAATACACACACAGTGTAGAGACCCCAAGCAAAACCAA 564
QY 564 ATACAGGATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTTGCTAAGGG 623
Db 565 ATACAGGATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTTGCTAAGGG 624
QY 624 TGTTCCTGCTGATAGTGTAGTATGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
Db 625 TGTTCCTGCTGATAGTGTAGTATGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 684
QY 684 AGTTCTCTGTTGAAGAACCAATACCATTTTCAAGCCCGCTGTTTCTAAGATTTATCCCTGA 743
Db 685 AGTTCTCTGTTGAAGAACCAATACCATTTTCAAGCCCGCTGTTTCTAAGATTTATCCCTGA 744
QY 744 GAAATGCAATATCAAGTCTTATCCACCAAGTGTGAATAACGCTTGCACAAAGCTGAGCA 803
Db 745 GAAATGCAATATCAAGTCTTATCCACCAAGTGTGAATAACGCTTGCACAAAGCTGAGCA 804
QY 804 GGTGACATGACAAACCTGATTTATAGTGAAGACCAAGACTGTTCCCAAGAGATAAG 863
Db 805 GGTGACATGACAAACCTGATTTATAGTGAAGACCAAGACTGTTCCCAAGAGATAAG 864
QY 864 CAGCCAGCTGAAACAGCCCAAGACTTATAGTGAAGACCAAGACTGTTCCCAAGAGATAAG 923
Db 865 CAGCCAGCTGAAACAGCCCAAGACTTATAGTGAAGACCAAGACTGTTCCCAAGAGATAAG 924
QY 924 GGAATAACTGGAAGCTGAAAGTCAAAAGGAGGAGAGAGAAAGAACTTGAATGATGTC 983
Db 925 GGAATAACTGGAAGCTGAAAGTCAAAAGGAGGAGAGAGAAAGAACTTGAATGATGTC 984
QY 984 TGCATTTGAACAAATTTCTCGTGACATCTACAGGATGCTGATGAAGATATGCGAGGGC 1043
Db 985 TGCATTTGAACAAATTTCTCGTGACATCTACAGGATGCTGATGAAGATATGCGAGGGC 1044
QY 1044 CATGATGAAGTCAATCGTGAATCAAAATGSCACTGTTCTCAACCAATTTGGAAGATGT 1103
Db 1045 CATGATGAAGTCAATCGTGAATCAAAATGSCACTGTTCTCAACCAATTTGGAAGATGT 1104
QY 1104 TGGAGCAAAAGAGTGAAGGAGGCCCTGATGATGAGTCAAGAGTGGGAATA 1163
Db 1105 TGGAGCAAAAGAGTGAAGGAGGCCCTGATGATGAGTCAAGAGTGGGAATA 1164
QY 1164 CTAAAGTTTGGAGTGGCCGCTTTTGTAAATCCAGGTCTTGGAACTATGACCTAATCT 1223
Db 1165 CTAAAGTTTGGAGTGGCCGCTTTTGTAAATCCAGGTCTTGGAACTATGACCTAATCT 1224
QY 1224 GCCCACCCTAGTGCCATGAGCTTGTGTTAAAGTCTGCTTTTAAAGCTTTTAAAGCTTTCTGTA 1283
Db 1225 GTCCACCCTAGTGCCATGAGCTTGTGTTAAAGTCTGCTTTTAAAGCTTTTAAAGCTTTCTGTA 1284
QY 1284 TGAAGTCTAGCTGTGTCTGAGTGTACTTGGCTAGCTTTTGAAGCATAGGATATGC 1343
Db 1285 TGAAGTCTAGCTGTGTCTGAGTGTACTTGGCTAGCTTTTGAAGCATAGGATATGC 1344
QY 1344 TATTATGTTGATGACCCGGAACCAATCTGTGTTATCAAGACAGCAAACTAATATGAT 1403
Db 1345 TATTATGTTGATGACCCGGAACCAATCTGTGTTATCAAGACAGCAAACTAATATGAT 1404
QY 1404 GGAATCCCTATTCTTTTGGTCTAAA 1430
Db 1405 GGAATCCCTATTCTTTTGGTCTAAA 1431
```

RESULT 3

US-10-425-114-25244

; Sequence 25244, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 25244

; LENGTH: 1450

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB732-024-C2_FLI

US-10-425-114-25244

Query Match

Best Local Similarity 95.8%; Score 1394.2; DB 17; Length 1450;

Matches 1399; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 24 AGTCATCGTCTGCTCGCCGCGCAGAGCGCAGACGCCCCAAAGCCGGAAGGGCGCCAT 83
Db 31 AGTCATCGTCTGCTCGCCGCGCAGAGCGCAGACGCCCCAAAGCCGGAAGGGCGCCAT 90
QY 84 GCGCGGTGCGATCTGGAGAGCAAGCCCAAGGAGGCTTCGTGAGAGGACATTCGAGCT 143
Db 91 GCGCGGTGCGATCTGGAGAGCAAGCCCAAGGAGGCTTCGTGAGAGGACATTCGAGCT 150
QY 144 GCGCACCGGAGCTCTACAGCAGGCGCATCGACGCGGCGCGCCGCGCCGCGCATCTATGC 203
Db 151 GCGCACCGGAGCTCTACAGCAGGCGCATCGACGCGGCGCGCCGCGCCGCGCATCTATGC 210
QY 204 CGACGCGCGCCGCGCGCAGCAGTCAAGCTCGGCAACTACACTGAGGCTGTGCGGATGTAA 263
Db 211 CGACGCGCGCCGCGCGCAGCAGTCAAGCTCGGCAACTACACTGAGGCTGTGCGGATGTAA 270
QY 264 CAAGCAATTTGAGCTTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 323
Db 271 CAAGCAATTTGAGCTTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 330
QY 324 CATTAAGCTTGAAGATACCAAACTGCAAGGCTGCTCTTGAAGTGGTCTTCTTCTTATGC 383
Db 331 CATTAAGCTTGAAGATACCAAACTGCAAGGCTGCTCTTGAAGTGGTCTTCTTCTTATGC 390
QY 384 ATCAGCGGATCAAGGTTTGTCTGCTATTTGAAGGAATGTGATGAGCGCATCGCTGAGGA 443
Db 391 ATCAGCGGATCAAGGTTTGTCTGCTATTTGAAGGAATGTGATGAGCGCATCGCTGAGGA 450
QY 444 ATCTAGCCAGGACCAAGTAAAGATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
Db 451 ATCTAGCCAGGACCAAGTAAAGATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 510
QY 504 GGAGGATTTTCAAAATATGAGAAATACACACACAGTGTAGAACCCCAAGCAAAACCAA 563
Db 511 GGAGGATTTTCAAAATATGAGAAATACACACACAGTGTAGAACCCCAAGCAAAACCAA 570
QY 564 ATACAGGATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTTGCTAAGGG 623
Db 571 ATACAGGATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTTGCTAAGGG 630
QY 624 TGTTCCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
Db 631 TGTTCCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 690
QY 684 AGTTCTCTGTTGAAGAACCAATACCATTTTCAAGCCCGCTGTTTCTAAGATTTATCCCTGA 743
Db 691 AGTTCTCTGTTGAAGAACCAATACCATTTTCAAGCCCGCTGTTTCTAAGATTTATCCCTGA 750
```


QY 1300 GTCTAGTGTCTACTTGGCGTAGCTTTTGTATGGACATA-----GGTATGCTATTATGTGGTA 1355
DB 1321 GTCTAGTGTCTTGTGCGTAGTTTTTGTATGGACATATCATGTATGCTATTCTGTGGTA 1380
QY 1356 TCGACCCCGAACCATCTGTGGTATCAAAAGACAG-CAAACTATATATGTATGGAATGCTAT 1414
DB 1381 TC-----ACATCTGTGTAAACAAAGACAGTAAACTATAATCTGTGGAATGCTAT 1432
QY 1415 TCCTTTTGTCTAAAAA 1431
DB 1433 TCCTTTTGTCTAAAGA 1449

RESULT 7

US-10-425-114-25845
; Sequence 25845, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25845
; LENGTH: 1361
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3960-013-B4_FLI
US-10-425-114-25845

Query Match 72.0%; Score 1047; DB 17; Length 1361;

Best Local Similarity 92.3%; Pred. No. 2.1e-285;

Matches 1125; Conservative 0; Mismatches 90; Indels 4; Gaps 2;

QY 54 GCAGAGCCCAAGCCGAGGGCGCCATGGCGCGTGGATCTGGAGAGCAAGGCCAA 113
DB 71 GCAGAGCCCGAGAGCGGAGGAGCCATGGCGCGTGGATCTGGAGAGCAAGGCCAA 130
QY 114 GGAGGCTTCTGTCGACGACGACTTCGAGCTGGCCACCAGCTCTACAGCCAGGCCCATCGA 173
DB 131 GGAGGCTTCTGTCGACGACGACTTCGAGCTGGCCGCCGAACTCTACAGCCAGGCCCATCGA 190
QY 174 CGCGGGCGCCGACCGCCGACCTCTATGCCGCGCCGCGCCGACATCAAGCTCGG 233
DB 191 CGCGGGCGCCGACCTGCGGATCTACGCGGACCGCGCCGCGCCGACATCAAGCTCGG 250
QY 234 CAATACATGAGGCTGTGGCGGATGCTAAAGAAATTGAGCTTGATCTATGATGCA 293
DB 251 CAATACATGAGGCTGTGGCAGATGCTAAAGAAATTGAGCTTGATCTATGATGCA 310
QY 294 TAAAGCTTACTACCGGAAAGTCTGTCATGCTTAAAGCTTGAAGAAATACCAACTGCAA 353
DB 311 TAAAGCTTACTACCGGAAAGTCTGTCATGCTTAAAGCTTGAAGAAATACCAACTGCAA 370
QY 354 GGCTGCTCTTGGTGTGGTTCTTCTATGATCATGAGCGATTCAAGGTTTGTCTGCTATT 413
DB 371 GGCTGCTCTTGGTGTGGTTCTTCTATGATCATGAGCGATTCAAGGTTTGTCTGCTATT 430
QY 414 GAAGGAATGTGATGAGCGCATGCTGAGGAATCTAGCCAGGACCCAGTAAAGAAATGTGA 473
DB 431 GAAGGAATGTGATGAGTGCATGCTGAGGAATCTAGCCAGGACCCAGTAAAGAAATGTGA 490
QY 474 GGCTACTGTGGTCTACTATTGAGGACAGGAGGATTTTCAAAATATGAGGAATACACC 533
DB 491 GGCTCTGTAGTGTCTACTGTGAGGACAGGAGGATGTCGCAAAATATGGAATATACACC 550

QY 534 ACCAGTGATAGACCCCGAGCAAAACCAAAATACAGGCATGACTACTACACAGTGCAC 593
DB 551 GCCAGTGGTAGAACCCCGAGCAAACTAAATATAGGCATGACTACTACACAGTGCAC 610
QY 594 AGAAGTGGTACTGACAAATATTTCTAAGGCTGTTCTTCTGCTGATAGTGTAGTCAATTGATT 653
DB 611 AGAAGTGGTCTCTACAAATATATGCTAAGGCTGTTCTTCTGCTGATAGTGTAGTCAATTGATT 670
QY 654 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGTTGAAGAACCAATACCAATTTTCA 713
DB 671 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGTTGAAGAACCAATACCAATTTTCA 730
QY 714 GCCCGTCTGTTTCTTAAGATTATCCCTGAGAAATGCAATATATCAAGTCTTATCCACAA 773
DB 731 GCCCGTCTGTTTCTTAAGATTATCCCTGAGAAATGCAAGTATCAAGTCTTATCCACAA 790
QY 774 GGTGGAATACGCTTTCGCAAAAGCTGAGCAGGTGACATGGACAACTCTGATATATAGTGG 833
DB 791 GGTGGAATACGCTTTCGCAAAAGCTGAGCAGGTGACATGGACAACTCTGATATATAGTGG 850
QY 834 AAGACCAAGACTGTTTCCCGAGAAATAGACGCGCAGCTGAAACAGCCCAAGACCTTC 893
DB 851 AAGACCAAGACTGTTTCCCGAGAAATAGACGCGCAGCTGAAACAGCCCAAGACCTTC 910
QY 894 ATACCCATCTTCAAGCGGAGAAAGAACTGGGATTAACCTGGAAGCTGAAGTCAAAAGGA 953
DB 911 ATACCCATCTTCAAGCTCAAAAGAAAGCTGGGATTAACCTGGAAGCTGAAGTCAAAAGGA 970
QY 954 GGAGAAGGAAGAAACTTGTGATGCTGCTGATTTGAAGTGTGCTGATTTGAACAGTCTTCCGATATCTA 1013
DB 971 GGAGAAGGAAGAAACTTGTGATGCTGCTGATTTGAAGTGTGCTGATTTGAACAGTCTTCCGATATCTA 1030
QY 1014 CAAGGATGCTGATGAAGATATGCGAGGCGCATGATGAAGTCAATTCGTGGAATCAATGG 1073
DB 1031 CAAGGATGCTGATGAAGATATGCGAGGCGCATGATGAAGTCAATTCAGGGAATCTAATGG 1090
QY 1074 CACTGTTCTCTCAACCAATTTGGAAGATGTTGAGCAAAAGAGTGAAGGAGGAGCCCC 1133
DB 1091 TACCGTTCTCTCAACCAATTTGGAAGATGTTGATCAAAAGACGGTGAAGAGGAGCCCTCC 1150
QY 1134 TGATGTTATGAGCTCAAGAGTGGGAATACTAAAGTTTGGACTGCCCGCTTTTGTAAA 1193
DB 1151 TGATGTTATGAGCTCAAGAGTGGGAATTTTGAAGTTTGGACTGCCCGCTTTTGTAAA 1208
QY 1194 TCAGGCTTTGGAAACTATGACCTAATTTCTGCCACCCCA--TAGTGCCATGAGCTTGCCTT 1251
DB 1209 TCCAGTCTTGGAGTTATGACCTAGTTCTGTCCACCCCATTTGTTGTTGTTGGTTTGGTT 1268
QY 1252 GGTAAAGTCTGCTTTTGG 1270
DB 1269 AAGTGTGTGTGTTTGG 1287

RESULT 8

US-10-425-114-22598
; Sequence 22598, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22598
; LENGTH: 1490

```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3591-005-E5_FLI
US-10-425-114-22598

Query Match      72.0%; Score 1047; DB 17; Length 1490;
Best Local Similarity 92.3%; Pred. No. 2.2e-285;
Matches 1125; Conservative 0; Mismatches 90; Indels 4; Gaps 2;

QY 54 GCAGAGCCCAAGCGGAGGCGCCATGCGCGGTCGGATCTGGAGAGCAAGGCCAA 113
DB 139 GCAGAGCCCAAGCGGAGGAGCGCCATGCGCGGTCGGATCTGGAGAGCAAGGCCAA 198
QY 114 GGAGGCTTTGTCGACGACGACTTCGAGCTGGCCACCGGACTCTACAGCGAGGCCATCGA 173
DB 199 GGAGGCTTTGTCGACGACGACTTCGAGCTGGCCACCGGACTCTACAGCGAGGCCATCGA 258
QY 174 CGCGGGCCCGCCACCGCGACCTCTATGCGGACCGCGCCAGCGGCACATCAAGCTCGG 233
DB 259 CGCGGGCCCGCCACCTGCGGATCTCTACGCGACCGCGCCAGCGGCACATCAAGCTCGG 318
QY 234 CAACTACACTGAGGCTGTGGCGGATGCTTAACAAAGCAATTGAGCTTGATCTTATGATGA 293
DB 319 CAACTACACTGAGGCTGTGGCAGATGCTTAACAAAGCAATTGAGCTTGATCTTATGATGA 378
QY 294 TAAAGCTTACTACCGAAGGTCGTCATGATTAAGCTTGAAGATACCAACTGCCAA 353
DB 379 TAAAGCTTACTACCGAAGGTCGTCATGATTAAGCTTGAAGATACCAACTGCCAA 438
QY 354 GGCTGCTCTGAGTGGGTTCTTCTATGATGAGGATTAAGCTTGAAGTGGTCTGCTGCTATT 413
DB 439 GGCTGCTCTGAGTGGGTTCTTCTATGATGAGGATTAAGCTTGAAGTGGTCTGCTGCTATT 498
QY 414 GAGGAGATGATGAGCGATCTGCTGAGGATCTAGCGGACCGGACCGGATTAAGCTTGA 473
DB 499 GAGGAGATGATGAGCGATCTGCTGAGGATCTAGCGGACCGGACCGGATTAAGCTTGA 558
QY 474 GGCTACTGTGGCTGCTACTATTGAGGACGAGGAGATTTACAAATATGAGAAATACACC 533
DB 559 GGCTCTGTAGCTGCTACTGTTGAGGACGAGGAGATGTCGAATATGATTAATACACC 618
QY 534 ACCAGTGATAGAACCCCGGAGCAAAACAAATACAGCGATGACTACTACAAAGTGGCCAC 593
DB 619 GCCAGTGATAGAACCCCGGAGCAAAACAAATATAGCGATGACTACTACAAAGTGGCCAC 678
QY 594 AGAAGTGGTACTCACAAATATTTCTAAGGCTGTTCTGCTGATGAGTGTGATGATGATT 653
DB 679 AGAAGTGGTCTCACAAATATATGCTAAGGCTGTTCTGCTGATGAGTGTGATGATGATT 738
QY 654 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGTTGAAGAACCAATACCATTTTCA 713
DB 739 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGTTGAAGAACCAATACCATTTTCA 798
QY 714 GCGCCGTGTTTCTAAGATATCCCTGAGAAATGCAAAATATCAAGTCTTATFCCACAA 773
DB 799 GCGCCGTGTTTCTAAGATATCCCTGAGAAATGCAAAATATCAAGTCTTATFCCACAA 858
QY 774 GGTGGAATACGCTTGCAGGCTGAGCAGGTCATGAGCAACCCCTGATTTATAGTGG 833
DB 859 GGTGGAATACGCTTGCAGGCTGAGCAGGTCATGAGCAACCCCTGATTTATAGTGG 918
QY 834 AAGACCAAGACTGTTTCCCGAGAGATTAAGCAGCGGATGAAACAGCCCGGACCAAGCTTTC 893
DB 919 AAGACCAAGACTGTTTCCCGAGAGATTAAGCAGCGGATGAAACAGCCCGGACCAAGCTTTC 978
QY 894 ATACCCATCTTAAAGCGGAAAGAAAGACTGGGATAACTGGAAAGTGAAGTCAAAAGGA 953
DB 979 ATACCCATCTTAAAGCGTCAAAAGAAAGACTGGGATAACTGGAAAGTGAAGTCAAAAGGA 1038
QY 954 GGAAGAGGAGAAAGAACTGATGCTGATGCTGATGAAACAAATCTTCCGTCGATCTA 1013
DB 1039 GGAAGAGGAGAAAGAACTGATGCTGATGCTGATGAAACAAAGTCTTCCGTCGATCTA 1098

RESULT 9
US-10-425-115-18041
; Sequence 18041, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18041
; LENGTH: 1519
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116456C.1
US-10-425-115-18041

Query Match      71.9%; Score 1046; DB 18; Length 1519;
Best Local Similarity 92.3%; Pred. No. 4.3e-285;
Matches 1124; Conservative 0; Mismatches 90; Indels 4; Gaps 2;

QY 54 GCAGAGCCCAAGCGGAGGCGCCATGCGCGGTCGGATCTGGAGAGCAAGGCCAA 113
DB 139 GCAGAGCCCAAGCGGAGGAGCGCCATGCGCGGTCGGATCTGGAGAGCAAGGCCAA 198
QY 114 GGAGGCTTCTGTCGACGACGACTTCGAGCTGGCCACCGGACTCTACAGCGAGGCCATCGA 173
DB 199 GGAGGCTTCTGTCGACGACGACTTCGAGCTGGCCACCGGACTCTACAGCGAGGCCATCGA 258
QY 174 CGCGGGCCCGCCACCGCGACCTCTATGCGGACCGCGCCAGCGGCACATCAAGCTCGG 233
DB 259 CGCGGGCCCGCCACCTGCGGATCTCTACGCGACCGCGCCAGCGGCACATCAAGCTCGG 318
QY 234 CAACTACACTGAGGCTGTGGCGGATGCTTAACAAAGCAATTGAGCTTGATCTTATGATGA 293
DB 319 CAACTACACTGAGGCTGTGGCAGATGCTTAACAAAGCAATTGAGCTTGATCTTATGATGA 378
QY 294 TAAAGCTTACTACCGAAGGTCGTCATGATTAAGCTTGAAGATACCAACTGCCAA 353
DB 379 TAAAGCTTACTACCGAAGGTCGTCATGATTAAGCTTGAAGATACCAACTGCCAA 438
QY 354 GGCTGCTCTGAGTGGGTTCTTCTATGATGAGGATTAAGCTTGAAGTGGTCTGCTGCTATT 413
DB 439 GGCTGCTCTGAGTGGGTTCTTCTATGATGAGGATTAAGCTTGAAGTGGTCTGCTGCTATT 498
```

```
QY 414 GAAGGATGATGATGAGCGCATCGCTGAGGAACTCTAGCCAGGCCACCAAGATGTTGA 473
Db 499 GAAGGATGATGATGAGCGCATCGCTGAGGAACTCTAGCCAGGCCACCAAGATGTTGA 558
QY 474 GGCTACTGTGGCTGCTACTATTGAGGCAAGAGGAGATTTCACAAATGAGAGATACACC 533
Db 559 GGCTCTGTAGCTGCTACTGTGAGGCAAGAGGAGATGTCGCAAAATGATGATACACC 618
QY 534 ACCAGTGATAGAACCCCGCCAGCAAAACCAAAATACAGGCATGACTACTACAAAGTGCCAC 593
Db 619 GCCAGTGGTGAACCCCGCCAGCAAAACCAAAATAGGCATGACTACTACAAAGTGCCAC 678
QY 594 AGAAGTGGTACTGACAAATATTGCTAAGGGTGTTCCTGCTGATAGTCTGATGATTT 653
Db 679 AGAAGTGGTCTCACAATATATGCTAAGGGTGTTCCTGCTGATAGTCTGATGATTT 738
QY 654 TGGTGAACAGATGTTGAGTGATATCCATTGGAAGTTTCCTGCTGAAAGCAACATACATTTCA 713
Db 739 TGGTGAACAGATGTTGAGTGATATCCATTGGAAGTTTCCTGCTGAAAGCAACATTTCA 798
QY 714 GCCCGTCTGTTTCTTAAGATTTATCCCTGAGAAATGCAAAATATCAAGTCTTTATCCCAA 773
Db 799 GCCCGTCTGTTTCTTAAGATTTATCCCTGAGAAATGCAAAATATCAAGTCTTTATCCCAA 858
QY 774 GATTGAAATACGCTTGCAGAAAGCTGAGCAGGTGACATGACCAACCTGGATTTATAGTG 833
Db 859 GGTGAAATACGCTTGCAGAAAGCTGAGCAGGTGACATGACCAACCTGGATTTATAGTG 918
QY 834 AAGACCAAGATGTTTCCCGAGAGATTAAGCAGCGCAGCTGAAACAGCCCGCAAGACTTC 893
Db 919 AAGACCAAGCTATTCCCGAGAGATTAAGCAGCGCAGCTGAAACAGCCCGCAAGACTTC 978
QY 894 ATACCCTCTTCAAGCGGAAAGAAAGACTGGGATAAATCTGGAAGTGAAGTCAAAAGGA 953
Db 979 ATACCCTCTTCAAGCGGAAAGAAAGACTGGGATAAATCTGGAAGTGAAGTCAAAAGGA 1038
QY 954 GGAAGAGGAGAAAGAAAGCTGATGGTGTGATGCTGATGAAAGTCTTCCGTCACATCTA 1013
Db 1039 GGAAGAGGAGAAAGAAAGCTGATGGTGTGATGCTGATGAAAGTCTTCCGTCACATCTA 1098
QY 1014 CAAGGATGCTGATGAAGATATGGGAGGGCCATGATGAAGTCAATCTGGAATCAAAAGG 1073
Db 1099 CAAGGATGCTGATGAAGATATGGGAGGGCCATGATGAAGTCAATCTGGAATCAAAAGG 1158
QY 1074 CACTGTTCTCTCAACCAATTTGAAAGATGTTGAGCAAAAGATGAAGGAGGAGCCGCC 1133
Db 1159 TACCGTTCTCTCAACCAATTTGAAAGATGTTGAGCAAAAGATGAAGGAGGAGCCGCC 1218
QY 1134 TGAATGATGGAGCTCAAGAGTGGGAATGATGATGAAAGTCAATCTGGAATCAAAAGG 1193
Db 1219 TGAATGATGGAGCTCAAGAGTGGGAATGATGATGAAAGTCAATCTGGAATCAAAAGG 1276
QY 1194 TCCAGTCTTGGAAACTATGACCTAAATTCGCCACCCCA--TAGTGCATGAGCTTGCTT 1251
Db 1277 TCCAGTCTTGGAAACTATGACCTAAATTCGCCACCCCA--TAGTGCATGAGCTTGCTT 1336
QY 1252 GGTAAAGTCTCTGCTTTT 1269
Db 1337 AAGTGTGTGTGTGTGT 1354
```

RESULT 10
US-10-739-930-4858/c
; Sequence 4858, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088

```
; SEQ ID NO 4858
; LENGTH: 2471
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(2471)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER396_1
; US-10-739-930-4858

Query Match 51.8%; Score 753; DB 18; Length 2471;
Best Local Similarity 79.3%; Pred. No. 5.6e-202;
Matches 938; Conservative 0; Mismatches 200; Indels 45; Gaps 2;

QY 61 CCCAAAGCCGAGAGGCGCCATGCGCGCTCGGATCTCGAGAGCAAGGCCAAGAGGCC 120
Db 2303 CTCGACGACAGACATGCGCGCGCGCGCGCTCGGATCTCGAGAGCAAGGCCAAGAGGCC 2244
QY 121 TTGCTCGACGACGACATTCGAGCTGGCGACCGAGCTCTACAGCCAGGCCATCGACGCCGG 180
Db 2243 TTGCTCGACGACGACATTCGAGCTGGCGCGCGCGCTCTACAGCCAGGCCATTCGAGGCCGG 2184
QY 181 CCGGCCACCGCCGACCTCTATGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 2183 CCGGCCACCGCGGAACTCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2124
QY 241 ACTGAGGCTGTGCGCGATGCTAAAGCAATTTGAGCTTGATCTCTATGATGCATAAGCT 300
Db 2123 ACTGAGGCTGTAGCTGATGCCAAGCAATTTGAGCTTGATCTCTCTGATGCATNAAGCA 2064
QY 301 TACTACCGGAAAGGCTGCTCATGCTAAAGCTTTGAAGAAATACCAAACTGCAAAAGGCTGT 360
Db 2063 TACTCTCGAAGGGCTCTGCTTGCATCAAGCTGGAGGAAATACCAAACTGCAAAAGGCTGT 2004
QY 361 CTTGAGTGGGTTCTTCTTATGCTATGCGGATCGGAGTTCAGGTTTCTGCTCTATTGAAGGAA 420
Db 2003 CTTGAAAGTGGGTTCTTCTTATGCTATGCTGAGTCTCGAGGTTTACTCGTCTGATGAAAGGAG 1944
QY 421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGACCAAGTAAAGAAATTTGAGGCTACT 480
Db 1943 TGTGATGATGCTATTGCTGAGGAGGCTAGCCAGGTCCAGTAAAGAAATGCGGCTGCGGCT 1884
QY 481 GTGGCT-----GCTACTATTGAG 498
Db 1883 GTTGCTTCAGTACATCTTTCGGGGCATCTTCGGGGCTACAACTGTGCTACTGAAGCT 1824
QY 499 GACCAAGGAGATTTCAAAATATGGAATATACCAACCAAGTATAGAACCCCAAGCAAA 558
Db 1823 GAGGACCCGAGTGGTGCNAATATGGAGAAATGCAAGCCAAAGTGAAGTGCCTCAAGCAAG 1764
QY 559 CCAAAATACAGGCGATGACTACTACAAAGTGCACAGAGAGTGGTACTGCAAAATATTGCT 618
Db 1763 CCCAAATACAGGCGATGACTACTACAAATCTCTACAGAGTGGTACTGACTATATTGCT 1704
QY 619 AAGGGTGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
Db 1703 AAGGGTGTTCAGCTGACAGCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1644
QY 679 ATTGAAGTTCCTGGTGAAGAACCATACCAATTTTCAGCCCGCTCTGTTTCTTAAGATTTATC 738
Db 1643 ATTGAATCTTCGGTGGAGAACCATACCAATTTTCAGCCCGCTCTGTTTCTTAAGATTTATC 1584
QY 739 CTTGAGAAATGCAAAATATCAAGTCTTATCCCAAGGTTGAAATAGCCCTTTCGCAAAAGCT 798
Db 1583 CCAGATAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1524
QY 799 GAGCAGGTGACATGGAACCAACCTCGGATTTATAGTGAAGAACCAAGACTGTTTCCCAAGAG 858
Db 1523 GAGCCAGTAACTGGACATCAATTTGATTTATGATTTATGATTTATGATTTATGATTTATGAT 1467
QY 859 ATAAGCAGCGCAGCTGAAACAGCCCGCAAGACCTTCTATCCCATCTTCTCAAAAGGCGAAGAA 918
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 00:00:36 ; Search time 5229.64 Seconds
(without alignments)
10590.311 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcgacagcgccac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1378.2	94.7	1646	3 AY103953	AY103953 Zea mays
2	719	49.4	727	6 CA401716	CA401716 EL01N0424
3	708.2	48.7	752	6 CD443320	CD443320 EL01N0424
4	699	48.0	1131	9 CL959360	CL959360 OGI00002
5	660.6	45.4	686	6 CF041238	CF041238 OGI00002
6	657.2	45.2	772	6 CA258233	CA258233 SCCCR300
7	652.4	44.8	668	2 AW076274	AW076274 614063D11
8	650.8	44.7	798	6 CA176210	CA176210 SCULST102
9	647	44.5	756	6 CA184150	CA184150 SCRFST314
10	643.4	44.2	659	2 AW090936	AW090936 614068G07
11	641	44.1	759	6 CA245820	CA245820 SCEZFL508
12	637.4	43.8	778	6 CA227695	CA227695 SCULFL301
13	631	43.4	666	6 CA215140	CA215140 SC5BAD112
14	628.8	43.2	681	2 AW076372	AW076372 614066C10
15	627	43.1	658	2 AW155786	AW155786 614096F11
16	623.2	42.8	648	1 AI603781	AI603781 486096G09
17	615.4	42.3	665	6 CA153735	CA153735 SCVPRZ203
18	608.2	41.8	678	6 CA177070	CA177070 SCVCSF105
19	601.8	41.4	725	6 CA094787	CA094787 SCCCL401
20	601	41.3	609	7 CK826852	CK826852 zmrsub1.0
21	600.2	41.3	609	2 AW017784	AW017784 614063D11
22	599.4	41.2	640	6 CF002911	CF002911 QBH17B05
23	596.4	41.0	622	6 CA229980	CA229980 SCJFFL3C0
24	595.8	40.9	599	2 AW331768	AW331768 614096F11

25	594.8	40.9	730	2 BE362001	BE362001 DGI_83_H0
26	591	40.6	591	2 AW018202	AW018202 614068G07
27	589.8	40.5	593	4 BG265701	BG265701 1000034F0
28	584.4	40.2	594	1 AI601072	AI601072 486096G09
29	583.4	40.1	596	2 AW042392	AW042392 614028B03
30	583.4	40.1	609	6 CA229238	CA229238 SCAGFL302
31	579.8	39.8	718	6 CA106435	CA106435 SCQHR101
32	576.8	39.6	702	6 CA262995	CA262995 SCPRLB202
33	576	39.6	588	5 BU499669	BU499669 946178B11
34	571.2	39.3	680	6 CA128103	CA128103 SCAGLR203
35	567.8	39.0	578	2 AW000264	AW000264 614009H02
36	565.2	38.8	594	6 CA098274	CA098274 SCMCCL605
37	563	38.7	574	4 BI64145	BI64145 949022C08
38	560.6	38.5	914	6 CA084222	CA084222 SC5QAM203
39	560.4	38.5	579	2 AW054458	AW054458 660007D01
40	560.2	38.5	607	1 AI601071	AI601071 486096G09
41	558.4	38.4	560	2 AW018065	AW018065 614066C10
42	557.2	38.3	630	6 CA486265	CA486265 WHE4329.D
43	556.4	38.2	578	6 CD996630	CD996630 QBC6a07.X
44	555.6	38.2	579	2 AW066518	AW066518 660015G08
45	555.6	38.2	736	6 CA227777	CA227777 SCJLFL301

ALIGNMENTS

RESULT 1
AY103953
LOCUS Zea mays PC0098564 mRNA sequence. 1646 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0098564 mRNA sequence.
ACCESSION AY103953
VERSION AY103953.1 GI:21207031
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1646)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1646)
Coe,S.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers
1..1646
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635643"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 94.7% Score 1378.2; DB 3; Length 1646;

D	b	1163	CATGATGAAGTCAATCGTGGAATCAAAAGGCACGGTCCTCACCACCATTTGGGGAATAA	
Q	y	1104	TGGAGCAAAGAAGGTAGAAGGGAGCCGCCCTGATGTTATGGAGCTCAAGAAGTGGGAATA	1163
D	b	1223	TGGAGCAAAGAAGGTAGAAGGGAGCCGCCCTGATGTTATGGAGCTCAAGAAGTGGGAATA	1282
Q	y	1164	CTAAAGTTTGGACTGCCCTCTTTTCTTAATCCAGGTCTTGGAAACTATGACCTAATTCT	1223
D	b	1283	CTAAAGTTTGGACTGCCCTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCT	1342
Q	y	1224	GCCCACCATAGTGCCATGAGCTTGGTTGGTTAAAGTCTCTGCTTTTGTAAAGCTTCTGTA	1283
D	b	1343	GTCCACCCATAGTGCCATGAGCTTGGTTGGTTAAAGTCTCTGCTTTTGTAAAGCTTCTGTA	1402
Q	y	1284	TGACAGTCTAGCGTGTCTGAGTGTCTACTTTCGGTAGCTTTTATGACGACATAGTATGC	1343
D	b	1403	TGACAGTCTAGCGTGTCTGAGTGTCTACTTTCGGTAGCTTTTATGACGACATAGTATGC	1462
Q	y	1344	TATTATGTGTGTATGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT	1403
D	b	1463	TATTATGTGTGTATGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT	1522
Q	y	1404	GGAATGCCTATCTTTTTTGGTCTAATAA	1430
D	b	1523	GGAATGCCTATCTTTTTTGGTCTAAGA	1549
RESULT 2				
CA401716/c				
LOCUS				
DEFINITION EL01N0424G10.g Endosperm_4 Zea mays cDNA, mRNA sequence.				
ACCESSION CA401716				
VERSION CA401716.1 GI:24766567				
KEYWORDS EST.				
SOURCE				
ORGANISM				
Zea mays				
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE				
AUTHORS Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F., Larkins,B., Beecraft,P. and Messing,J.				
TITLE Characterization of the maize endosperm transcriptome and its comparison to the rice genome				
JOURNAL Genome Res. 14 (10), 1932-1937 (2004)				
COMMENT Contact: Lai, Jinsheng Dr. Joachim Messing's lab Waksman Institute, Rutgers University 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA Tel.: 732-445-3801 Fax: 732-445-5735 Email: jlai@waksman.rutgers.edu Seq primer: T7. 				
FEATURES				
source				
Location/Qualifiers				
1..727				
/organism="Zea mays"				
/mol_type="mRNA"				
/cultivar="W22"				
(db_xref="taxon:4577")				
/tissue_type="Endosperm of 7-23DAP"				
/clone_lib="Endosperm_4"				
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"				
ORIGIN				
Query Match 49.4%; Score 719; DB 6; Length 727;				
Best Local Similarity 99.3%; Pred. No. 1.8e-175;				
Matches 722; Conservative 0; Mismatches 5; Indels 0; Gaps 0				
Q	y	712	CAGCCCCGCTGTTTCTAAGATTATCCCTGAGAATGCAATATCAAGCTTATCCACC	771
D	b	727	CAGCCCCGCTGTTTCTAAGATTATCCCTGAGAATGCAATATCAAGCTTATCCACC	668

772	Qy		AAGTTGAAATACGCCTTGC	AAAGCTGAGCAGGTGACATGGACAA	CCCTGGATTATAGT	831
773	Db					
667	Db		AAGTTGAAATACGCCTTGC	AAAGCTGAGCAGGTGACATGGACAA	CCCTGGATTATAGT	608
668	Db					
832	Qy		GGAGACCAAGACTGTTCC	CCAGAGATGAACGCCAGCTGAAACAG	CCCCCAAGCCT	891
833	Db					
607	Db		GGAGACCAAGACTGTTCC	CCAGAGATGAACGCCAGCTGAAACAG	CCCCCAAGCCT	548
608	Db					
892	Qy		TCATACCCATCTTCAAAG	CGCAAAAAGACTGGGATAAACTGGAAG	CTGAAGTCAAAAAG	951
893	Db					
547	Db		TCATACCCATCTTCAAAG	CGCAAAAAGACTGGGATAAACTGGAAG	CTGAAGTCAAAAAG	488
548	Db					
952	Qy		GAGGAGAGGAAGAAAACT	TGATGGTGATGCTGCATTTGAACAAAT	TTCTTCCGTGACATC	1011
953	Db					
487	Db		GAGGAGAGGAAGAAAACT	TGATGGTGATGCTGCATTTGAACAAAT	TTCTTCCGTGACATC	428
488	Db					
1012	Qy		TACAAGGATGCTGATGAAG	ATATCGGAGGGCCATGATGAAGTCA	TTCTGTTGGAAATCAAAT	1071
1013	Db					
427	Db		TACAAGGATGCTGATGAAG	ATATCGGAGGGCCATGATGAAGTCA	TTCTGTTGGAAATCAAAT	368
428	Db					
1072	Qy		GGCACTGTTCTCTCAACC	CAATTGGGAAGATGTTGGACAAAGAGG	TAGGAAGGGAGCCCC	1131
1073	Db					
367	Db		GGCACTGTTCTCTCAACC	CAATTGGGAAGATGTTGGACAAAGAGG	TAGGAAGGGAGCCCC	308
368	Db					
1132	Qy		CCTGATGGTATGGAGCT	CAAGAAAGTGGGAATACATAAAGTTT	TGGACTGCCCGTCTTTTGTA	1191
1133	Db					
307	Db		CCTGATGGTATGGAGCT	CAAGAAAGTGGGAATACATAAAGTTT	TGGACTGCCCGTCTTTTGTA	248
308	Db					
1192	Qy		AATCCAGGTCTTGGAACT	ATGACCTAAATTTCTGCCCA	CCCATAGTGCCATGAGCTTGCTT	1251
1193	Db					
247	Db		AATCCAGGTCTTGGAACT	ATGACCTAAATTTCTGCCCA	CCCATAGTGCCATGAGCTTGCTT	188
248	Db					
1252	Qy		GGTTAAGTCTCTGCTTTT	GTAAGCTTCTGTATGACAGTCTAGG	GTGTGCTGAGTGCTA	1311
1253	Db					
187	Db		GGTTAAGTCTCTGCTTTT	GTAAGCTTCTGTATGACAGTCTAGG	GTGTGCTGAGTGCTA	128
188	Db					
1312	Qy		CTTGGCGTAGCTTTT	TGATGGACATAGTATGCTATTATG	TGGTATCGACCCCGAACCATC	1371
1313	Db					
127	Db		CTTGGCGTAGCTTTT	TGATGGACATAGTATGCTATTATG	TGGTATCGACCCCGAACCATC	68
128	Db					
1372	Qy		TGTGGTATCAAAAGACAG	CAAACTATATGATGGAAATGCC	TATTTCTTTTGGTCTAAAAA	1431
1373	Db					
67	Db		TGTGGTATCAAAAGACAG	CAAACTATATGATGGAAATGCC	TATTTCTTTTAAAAA	8
68	Db					
1432	Qy		AAAAAA	1438		
1433	Db					
7	Db		AAAAAA	1		

RESULT	3
CD443320	
LOCUS	CD443320
DEFINITION	EL01N0424G10.b Endosperm
	752 bp mRNA linear EST 03-JUN-2003
	4 Zea mays cDNA, mRNA sequence.

ACCESSION NUMBER	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL	COMMENT
CD443320			GI:31358963					
CD443320.1			EST.					
			Zea mays					
			Zea mays					
			Eukaryota					
			Viridiplantae					
			Streptophyta					
			Embryophyta					
			Tracheophyta					
			Spermatophyta					
			Magnoliophyta					
			Liliopsida					
			Poales					
			Poaceae					
			PACCAD					
			Clade					
			Panicoidae					
			Andropogoneae					
			Zea					
			1 (bases 1 to 752)					
			Lai, J.,					
			Dey, N.,					
			Kim, C.S.,					
			Bharti, A.K.,					
			Rudd, S.,					
			Mayer, K.F.,					
			Larkins, B.,					
			Becraft, P. and					
			Messing, J.					
			Characterization of the maize endosperm transcriptome and its comparison to the rice genome					
			Genome Res. 14 (10), 1932-1937 (2004)					
			Contact: Lai, Jinsheng					
			Dr. Joachim Messing's lab					
			Waksman Institute, Rutgers University					
			190 Frelinghuysen Rd., Piscataway, NJ 08854, USA					
			Tel: 732-445-3801					
			Fax: 732-445-5735					

```
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
FEATURES
    source
        1..752
            /organism="Zea mays"
            /mol_type="mRNA"
            /cultivar="W22"
            /db_xref="taxon:4577"
            /tissue_type="Endosperm of 7-23DAP"
            /clone_lib="Endosperm 4"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                XhoI"
```

ORIGIN	Query Match	48.7%;	Score 708.2;	DB 6;	Length 752;
	Best Local Similarity	99.6%;	Pred. No. 1.2e-172;		
	Matches 710;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	24	AGTCATCGTCTGGTCTCGTCCCGCAGAGCGCAGACGCCCAAGCCGGAAGGGCGGCCCAT	83		
DB	40	AGTCATCGTCTGGTCTCGTCCCGCAGAGCGCAGACGCCCAAGCCGGAAGGGCGGCCCAT	99		
QY	84	GGCCGCGTCCGATCTGGAGAGCAAGGCCAAGAGGCGCTTCGTCGACGACGACTTCGAGCT	143		
DB	100	GGCCGCGTCCGATCTGGAGAGCAAGGCCAAGAGGCGCTTCGTCGACGACGACTTCGAGCT	159		
QY	144	GGCCACCGAGCTTACAGCAGCGGCCCATCGACGCGGGCCGCGCACCGCGCAGCCTCTATGC	203		
DB	160	GGCCACCGAGCTTACAGCAGCGGCCCATCGACGCGGGCCGCGCACCGCGCAGCCTCTATGC	219		
QY	204	CGACCGCGCCGAGGCGGCACATCAAGCTCGGCACTACACTGAGGCTGCGCGGATGCTAA	263		
DB	220	CGACCGCGCCGAGGCGGCACATCAAGCTCGGCACTACACTGAGGCTGCGCGGATGCTAA	279		
QY	264	CAAAAGCAATTGAGCTTGATCTTATGATGATATAAAGCTTACTACCGGAAAGGTGCTGCATG	323		
DB	280	CAAAAGCAATTGAGCTTGATCTTATGATGATATAAAGCTTACTACCGGAAAGGTGCTGCATG	339		
QY	324	CATTAAAGCTTGAAGAAATACCAAACTGCAAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC	383		
DB	340	CATTAAAGCTTGAAGAAATACCAAACTGCAAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC	399		
QY	384	ATCAGCGGATTCAGAGTTTGGCTCGTCTATGTAAAGGAATGTGATGAGCGCATCGCTGAGGA	443		
DB	400	ATCAGCGGATTCAGAGTTTGGCTCGTCTATGTAAAGGAATGTGATGAGCGCATCGCTGAGGA	459		
QY	444	ATCTAGCCAGGCACCAAGTAAAGAAATGTTGAGGCTACTGCGGCTGCTACTATTGAGGACAA	503		
DB	460	ATCTAGCCAGGCACCAAGTAAAGAAATGTTGAGGCTACTGCGGCTGCTACTATTGAGGACAA	519		
QY	504	GGAGGATTTTCAAAATATGGAGAAATACACCAACAGTGATAGAACCCCAAGCAAAACCAA	563		
DB	520	GGAGGATTTTCAAAATATGGAGAAATACACCAACAGTGATAGAACCCCAAGCAAAACCAA	579		
QY	564	ATACAGGCATGACTACTACAAACAGTGGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG	623		
DB	580	ATACAGGCATGACTACTACAAACAGTGGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG	639		
QY	624	TGTTCTCTGTAGTGTAGTCAATTCATTTGGTGGAACAGATGTTGAGTGATATCCATTGA	683		
DB	640	TGTTCTCTGTAGTGTAGTCAATTCATTTGGTGGAACAGATGTTGAGTGATATCCATTGA	699		
QY	684	AGTTCTCTGTGAAAGAACCATACATTTTCAGGCCCGTCTGTTTTCTTAAGATTA	736		
DB	700	AGTTCTCTGTGAAAGAACCATACATTTTCAGGCCCGTCTGTTTTCTTAAGATTA	752		

RESULT 4
CL959360
LOCUS
DEFINITION
Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
CL959360
ACCESSION

```

VERSION CL959360.1 GI:52373436
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Enarthroideae; Oryzoideae; Oryza.
REFERENCE 1 (bases 1 to 1131)
AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatics
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
    source          Location/Qualifiers
        1..1131
            /organism="Oryza sativa (indica cultivar-group)"
            /mol_type="genomic DNA"
            /db_xref="taxon:39946"
            /clone_lib="Oryza sativa Express Library"
            /note="Oryza sativa exon trapped genomic sequences"
ORIGIN
Query Match          48.0%; Score 699; DB 9; Length 1131;
Best Local Similarity 78.6%; Pred. No. 3.3e-170;
Matches 894; Conservative 0; Mismatches 195; Indels 48; Gaps 3;
QY 73 AGGGGGCCGATGCGCGCTGCGATCTGGAGAGCAAGGCCAAGAGGCGCTTCGTCGACGAC 132
DB 1 ATGGCAACCGCCGCGCGCTGCGATCTGGAGAGCAAGGCCAAGGCGCGCTTCGTCGACGAC 60
QY 133 GACTTCGAGTGGCCACCGAGCTCTACAGCCAGCCATCGAGCGCGCGCCGCCACCGCC 192
DB 61 GACTTCGAGTGGCCACCGAGCTCTACAGCCAGCCATCGAGCGCGCGCCGCCACCGCC 120
QY 193 GACTTCATGCGACCGCGCCACCGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTG 252
DB 121 GAGCTCTACCGCACCGCGCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTA 180
QY 253 GCGGATGCTAACCAAGCAATTGAGCTTGATCTCTATGATGATCAATCAAGCTTACTACCGGAAA 312
DB 181 GCTGATGCTAACCAAGGCCATTGAATCTGACCCATCAATGCAAGGCTTATCTTCGTAAA 240
QY 313 GGTGCTGCAATGCAATTAAGCTTGAAGATACCAAGCTGCAAGGCTCTCTGAGTTGGGT 372
DB 241 GGGCTGCAATGATACGATCTGAGGAGTATCAACTGCAAGCGCTCTGGAATTTGGT 300
QY 373 TCTTCTTATGCAATCAGCGATTCAGGTTTGGCTGCTTATGAGGAATGTGATGAGCGC 432
DB 301 TACTCGTTCGATCTCTGCTGACTCAAGGTTTACTCGCTTAATGAAGGAGTGTGATGAGCGC 360
QY 433 ATGCTGAGGATCTAGCAGGACCGACCAAGTAAGATTTGA-----GGCT 477
DB 361 ATTGCTGAGGAGCTTATGAGGTCCTGTTAAGAGGCTGAAGATGAGCAGCTGCGGCC 420
QY 478 ACTGTGGCTGCTACTATTGAGGACAGGAGGATTTTCACAAATATGAGGAATACACACCA 537
DB 421 TCTGTTGCTCTTTTGTGAGGAAGGATGATGCTGCAACATGATATACACACCA 480
QY 538 GTGATAGAACCCCAAGCAAAACCAAAATACAGCATGACTTACAAACAGTGCCACAGAA 597
DB 481 ATGGTAGA-----AGTGAAGCCAAAATACAGGACGAGCTTCTTACAAAGTGCTACAGAA 534
QY 598 GTGGTACTGACAAATATTTGCTAAGGGTGTCTCTGCTGATAGTGTGATGATTTTGGT 657

```

```

535 GTTGATTGACAAATTTTGCAGAAAGGTTCTCTGTCGAGAAATGTTGTTGTTGATTGGT 594
658 GAACAGATGTTGAGTGTATCCATTGAAGTTCTCTGTTGAAGAACCAATACCAATTTTCAGCCC 717
595 GAACAAATGTTAAGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTCACCAATTTTCAGCT 654
718 CTTCTGTTTTCTAAGATTTATCCCTGAGAAATCAAAATATCAAGTCTTATCCACCAAGTT 777
655 CTTCTGTTTTCTAAGATTTATCCCTGAGAAATCAAAATATCAAGTCTTATCCACCAAGTT 714
778 GAATACGCTTGCAGAAAGCTGAGCAGGTGATGAGCAACCTGGATTTATGTAAGA 837
715 GAATAAGACTGCTGCTAAAGCTGAACAGATTACATGCACTCTGATTATGATAAAAA 774
838 CAAAGACTGTTTCCCGAGAGGATAAGCAGC-----CA 870
775 CCAAGGCTGTTCCCAAAAGATATCCCTCAGTTTACTGATAGCCCTTTCTTTATCA 834
871 GCTGAAACAGCCCCAAGACCTTTATACCCATCTTCAAGGCCGAAAAAGACTGGGATAAA 930
835 GCTGAATCGGCCAGAGGCCATCATATCTTCTCAAAATCCAAGAAAGACTGGGATAAA 894
931 CTGGAAGCTGAAGTCAAAAGGAGGAGAGGAGGAGAACTTGTGTTGATGCTGCTGATG 990
895 CTGGAAGCTGAAGTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954
991 AACAAATTTCTTCGTTGACATCTACAGGATGCTGATGAAGATATGCGGAGGCGCATGATG 1050
955 AACAAATTTCTTCGTTGACATCTACAGTGTGATGATGAAGATATGCGGAGGAGGAGGAGG 1014
1051 AAGTCTATTCGTTGAAATCAAAATGCGCTGTTCTCTCAACCAATGGAAGATGTTGAGCA 1110
1015 AAATCTTTTGTGTAATCTAACGTTACTGTTCTGTCGACCAATGGAAGATGTTGCTCG 1074
1111 AAGAGGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1167
1075 AAGAAGGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131

```

CF041238 686 bp mRNA linear EST 17-JUL-2003
 QC123f02.yg QCI Zea mays cDNA clone QC123f02, mRNA sequence.
 CF041238
 CF041238.1 GI:32936419
 EST.
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 Genoplante.
 1 (bases 1 to 686)
 Genoplante, a major partnership french program in plant genomics
 Unpublished (2003)
 Contact: Genoplante
 Genoplante
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>
 and <http://genoplante-info.infobiogen.fr/>).

FEATURES
 Location/Qualifiers
 1..686
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="F2"
 /db_xref="taxon:4577"
 /clone="QC123f02"
 /tissue_type="embryo"
 /clone_lib="QCI"

ORIGIN		COMMENT	
Query Match		Contact: Arruda P	
Best Local Similarity		Centro de Biologia Molecular e Engenharia Genetica	
Matches 674; Conservative 0; Mismatches 4; Indels 1; Gaps 1;		Universidade Estadual de Campinas	
		Caixa Postal 6010, 13083-970, Campinas SP, Brazil	
		Tel: 55 19 3788 1137	
		Fax: 55 19 3788 1089	
		Email: parruda@unicamp.br	
		Clone distribution: clone distribution information can be found	
		through the Brazilian Clone Collection Center (BCCC) at	
		http://www.bccccenter.fcav.unesp.br	
		Plate: 004 row: D column: 05	
		Seq primer: T7 Promoter Primer.	
		Location/Qualifiers	
FEATURES		1..772	
source		/organism="Saccharum officinarum"	
		/mol_type="mRNA"	
		/db_xref="taxon:4547"	
		/clone="SCCCT3004D05"	
		/lab_host="DH10B"	
		/clone_lib="RT3"	
		/notes="Organ: Root apex from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Root apex from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CU-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucet.lad.ic.unicamp.br/public"	
ORIGIN		Query Match 45.2%; Score 657.2; DB 6; Length 772;	
		Best Local Similarity 93.7%; Pred. No. 2.1e-159;	
		Matches 697; Conservative 0; Mismatches 43; Indels 4; Gaps 1;	
QY	18	CACGAGAGTCATCGTCTGCTGCGCGCGCAGAGCGCAGACCCCAAGCCGGAAGGGG	77
DB	33	CGCCCGAGTCATCGTCTGCTGCGCGCGCAGAGCGCAGACCCCAAGCC	88
QY	78	CGCCATGCGCGCGCTCGGATCTGGAGAGCAAGGCAAGGAGGCGCTTCGTCACGACGACTT	137
DB	89	CATGCGCGCGCGCTCGGATCTGGAGAGCAAGGCAAGGAGGCGCTTCGTCACGACGACTT	148
QY	138	CGAGCTGGCCACCGAGCTCTACAGCCAGGCGCATCGACGCGCGCGCGCCACCGCCGACCT	197
DB	149	CGAGCTGGCGCGCGAGCTCTACACCCAGGCGCATCGACGCGCGCGCGCCACCGCCGACCT	208
QY	198	CTATGCGGACCG	257
DB	209	CTAGCG	268
QY	258	TGCTAAACAAAGCAATTGAGCTTCGATCTATGATGATGATGATGATGATGATGATGATGATG	317
DB	269	TGCTAAACAAAGCAATTGAGCTTCGATCTATGATGATGATGATGATGATGATGATGATGATG	328
QY	318	TGCATGCATTAAAGCTTGAAGAATACCAAACTGCAAAAGGCTGCTCTTGAAGTTGGTTCTTC	377
DB	329	TGCATGCATTAAAGCTTGAAGAATACCAAACTGCAAAAGGCTGCTCTTGAAGTTGGTTCTTC	388
QY	378	TTATGATCAGCGGATTCAGGTTTGTCTGTCTTATTTGAAGGAATGTGATGAGCGCATCGC	437
DB	389	CTATGATCAGCGGATTCAGGTTTGTCTGTCTTATTTGAAGGAATGTGATGAGCGCATCGC	448
QY	438	TGAGGAATCTAGCG	497
DB	449	AGAGGAATCTAGCG	508
QY	498	GGACAAAGGAGGATTTTCAAAAATATGGAGATAACCAACCAAGTATGATGATGATGATGATG	557
DB	509	GGACAAAGGAGGATTTTCAAAAATATGGAGATAACCAACCAAGTATGATGATGATGATGATG	568
QY	558	ACCAAAATACAGGCGATGACTACTACAAACAGTGCACAGAGTGTGATGATGATGATGATGATG	617


```

QY 385 TCAGGCGATTCAAGGTTGCTGCTCTATTTGAAGGAATGTGATGAGCCATCGCTGAGGAA 444
Db 301 TCAGGCGATTCAAGGTTTACTCGTCTATTTGAAGGAATGTGATGAGCCATCGCTGAGGAA 360
QY 445 TCTAGCAGGACAGGTAAGAAATGTTGAGGCTACTGCTGCTGCTACTATGAGGACAAG 504
Db 361 TCTAGCAGGACAGGTAAGAAATGTTGAGGCTACTGCTGCTGCTACTATGAGGACAAG 420
QY 505 GAGGATTTCAAAATATGGAGATATACACCAAGGATAGAAACCCCAAGCAACCAAAA 564
Db 421 GAGGATGCCACAAATATGGAGATATACACCAAGGATAGAAACCCCAAGCAACCTTAA 480
QY 565 TACAGGATGACTACTACACAGTGCCACAGAGTGCTACTGACAAATATTTGCTAAGGGT 624
Db 481 TACAGGATGACTACTACACAGTGCCACAGAGTGCTACTGACAAATATTTGCTAAGGGT 540
QY 625 GTTCTCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCATTTGAA 684
Db 541 GTTCTCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCATTTGAA 600
QY 685 GTTCTCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCATTTGAA 744
Db 601 GTTCTCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCATTTGAA 660
QY 745 AATGCAATATCAAGTCTTATCCCAAGGTTGAATAGCGCTTGCAAAAGCTGAGCAG 804
Db 661 AATG-ANATATCAAGTCTTATCCCAAGGTTGAATAGCGCTTGCAAAAGCTGAGCAG 717
QY 805 GTGACATGGACAAACCTTGAATATAGTGGAGACCAAG 843
Db 718 GTGACATGGACAAACCTTGGGA-TATAGTGGAGACCAAG 755

```

```

RESULT 10
AW090936 659 bp mRNA linear EST 18-OCT-1999
LOCUS 614088G07.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW090936
VERSION AW090936.1 GI:6056546
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 659)
Walbot.V
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
955 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614068 row: G column: 07.
Location/Qualifiers
1. .659
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOIR"
/clone_lib="614 - root cDNA library from Walbot Lab"
note="Organ: root; Vector: pBluescriptII SK+; Site 1:
EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"

```

```

FEATURES
source

```

ORIGIN

```

Query Match 44.2%; Score 643.4; DB 2; Length 659;
Best Local Similarity 99.8%; Pred. No. 7.9e-156;
Matches .644; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 321 ATCCATTAAAGCTTTGAAGAAATACCAAACTGCCAAAGGCTGCTCTTGGAGTGGGTCTTCTTA 380
Db 15 ATCCATTAAAGCTTTGAAGAAATACCAAACTGCCAAAGGCTGCTCTTGGAGTGGGTCTTCTTA 74
QY 381 TGCATCAGGCGATTCAAGGTTTCTGCTGCTATTTGAAGGAATGTGATGAGCGCATCGCTGA 440
Db 75 TGCATCAGGCGATTCAAGGTTTCTGCTGCTATTTGAAGGAATGTGATGAGCGCATCGCTGA 134
QY 441 GGAATCTAGCCAGCCAGTAAAGAAATGTTGAGGCTACTGCTGCTGCTACTATTGAGGA 500
Db 135 GGAATCTAGCCAGCCAGTAAAGAAATGTTGAGGCTACTGCTGCTGCTACTATTGAGGA 194
QY 501 CAAAGGAGGATTTCAAAATATGGAGAAATACCAACCAAGTGTATGAAACCCCAAGCAAAACC 560
Db 195 CAAAGGAGGATTTCAAAATATGGAGAAATACCAACCAAGTGTATGAAACCCCAAGCAAAACC 254
QY 561 AANAATCAGGCATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTTGCTAA 620
Db 255 AANAATCAGGCATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTTGCTAA 314
QY 621 GGGTGTCTCTGCTGATGATGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCAT 680
Db 315 GGGTGTCTCTGCTGATGATGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCAT 374
QY 681 TGAAGTTCCTGCTGATGATGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCAT 740
Db 375 TGAAGTTCCTGCTGATGATGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCAT 434
QY 741 TGAAGTTCCTGCTGATGATGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCAT 800
Db 435 TGAAGTTCCTGCTGATGATGTAGTCAATTTTGGTGAACAGATGTTGAGTGTATCCAT 494
QY 801 GCAGGTGACATGACAAACCTGATATATAGTGAAGACCAAGACCTGTTCCCAAGAGAT 860
Db 495 GCAGGTGACATGACAAACCTGATATATAGTGAAGACCAAGACCTGTTCCCAAGAGAT 554
QY 861 AAGCAGCCAGCTGAAAACAGCCCAAGACCTTATACCATCTTTCAAAGCGGCAAAAAGA 920
Db 555 AAGCAGCCAGCTGAAAACAGCCCAAGACCTTATACCATCTTTCAAAGCGGCAAAAAGA 614
QY 921 CTGGGATTAACCTGGAGCTGAAAGTCAAAAAGGAGGAGGAGGA 965
Db 615 CTGGGATTAACCTGGAGCTGAAAGTCAAAAAGGAGGAGGAGGA 659

```

RESULT 11

```

CA245820 759 bp mRNA linear EST 25-SEP-2003
LOCUS SCEZFL5084D06.g Saccharum officinarum FL5 Saccharum officinarum
DEFINITION cDNA clone SCEZFL5084D06 5', mRNA sequence.
ACCESSION CA245820
VERSION CA245820.1 GI:35324833
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 759)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil

```


Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>

Plate: 084 row: D column: 06

Seq primer: T7 Promoter Primer.

Location/Qualifiers

FEATURES

1. .759
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEZPL5084D06"
/lab_host="DH10B"

/clone_lib="Saccharum officinarum FL5"
/notes="Organ: Developed inflorescence (20cm-long) without rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Developed inflorescence (20cm-long) without rachis]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 44.1%; Score 641; DB 6; Length 759;

Best Local Similarity 93.6%; Pred. No. 3.5e-155;

Matches 700; Conservative 0; Mismatches 43; Indels 5; Gaps 3;

QY 288 GATGCATAAGCTTACTACCGAAGAGTGCTGCATTAAGCTTGAAGATACCAAC 347

Db 2 GATGCATAAGCTTACTACCGAAGAGTGCTGCATTAAGCTTGAAGATACCAAC 61

QY 348 TGCAGAGGCTCTTGGAGTGGTCTTCTTATGATCAGGCGATTCAAGTTTCTCG 407

Db 62 TGCAGAGGCTCTTGGAGTGGTCTTCTTATGATCAGGCGATTCAAGTTTCTCG 121

QY 408 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCAGGCACGATGAAGAA 467

Db 122 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCAGGCACGATGAAGAA 181

QY 468 TGTGAGGCTACTGTGGCTGCTACTATTGAGGCAAGAGGATTTCAAAATATGAGAA 527

Db 182 GGTGAGGCTCTGTGGCTGCTACTATTGAGGCAAGAGGATTTCAAAATATGAGAA 241

QY 528 TACACACGATGATAGAACCCCAAGCAACCAAAATACAGGCATGACTACTACACAG 587

Db 242 TACACCGCCAGTGGTGAACCCCAAGCAACCAAAATACAGGCATGACTACTACACAG 301

QY 588 TGCACAGAGTGGTACTGACAAATATTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCA 647

Db 302 TGCACAGAGTGGTGTGACAAATATTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCA 361

QY 648 TGATTTGGTGAACAGATGTGAGTGTATCCATTGAAGTTTCTGGTGAAGAACCATACCA 707

Db 362 TGATTTGGTGAACAGATGTGAGTGTATCCATTGAAGTTTCTGGTGAAGAACCATACCA 421

QY 708 TTTTCAGCCCGCTGTGTTTCTAAGATTATCCCTGAGAAATGCAATATCAAGTCTTATC 767

Db 422 TTTTCAGCCCGCTGTGTTTCTAAGATTATCCCTGAGAAATGCAATATCAAGTCTTATC 481

QY 768 CACCAAGTTGAATACGCCCTTGCACAAAGCTGAGCAGGTGACATGACACCCCTGGATTA 827

Db 482 CACCAAGTTGAATACGCCCTTGCACAAAGCTGAGCAGGTGACATGACACCCCTGGATTA 541

QY 828 TAGTGAAGACCAAGACTGTCTCCCAAGAGATGAAGCAGCCAGCTGAACACGCCCAAG 887

Db 542 TAGTGAAGACCAAGACTGTCTCCCAAGAGATGAAGCAGCCAGCTGAACACGCCCAAG 601

QY 888 ACCTTCATACCCATCTTCAAAGCGGAAAAAGACTGG---GATAAACTGGAAGCTGAAGT 944

Db 602 ACCTTCATACCCATCTTCAAAGCGGAAAAAGACTGGGNGATAACTGGGAAGCTGATC 661

QY 945 CAAAAGAGGAGGAGAA--GGAAGAAAAAATTGATGGTGCATGCTGATGAACAAATTTTCC 1003

Db 662 AAAAGGGAAGAGGAGGAGAAACTTGTGATGCTTGTGATTTGNNACAATTC--TCC 720

QY 1004 GTGACATCTACAGGATGCTGATGAAGA 1031

Db 721 GTGACATCTACAGGATGCTGATGAAGA 748

RESULT 12

CA227695

LOCUS

DEFINITION

SCJLFL3014D05.g Saccharum officinarum FL3 Saccharum officinarum

cDNA clone SCJLFL3014D05 5', mRNA sequence.

CA227695

ACCESSION

CA227695.1 GI:35288469

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta;

Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

REFERENCE

1 (bases 1 to 778)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bcccenter.fcav.unesp.br>

Plate: 014 row: D column: 05

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. .778

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCJLFL3014D05"

/lab_host="DH10B"

/clone_lib="Saccharum officinarum FL3"

/note="Organ: Base of developing inflorescence (5cm-long);

Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [Base of

developing inflorescence (5cm-long)]. cDNA was prepared

from polyA+ mRNA using SuperScript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were

fractionated in a sepharose CL-2B 40cm-columns and

fragments sizing between 0.8 and 1.5 Kb were

directionally cloned into the vector. Details of each

source of RNA and library construction can be obtained at

<http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match

Best Local Similarity 43.8%; Score 637.4; DB 6; Length 778;

Matches 718; Conservative 0; Mismatches 56; Indels 7; Gaps 4;

QY 650 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAAGTTCTGTGTGAAGAACCATACCAT 709

Db 1 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAAGTTCTGTGTGAAGAACCATACCAT 60

QY 710 TTCAGCCCCGTCTGTTTCTTAAAGATTATCCCTGAGAAATGCAATATATCAAGTCTTATCCA 769

Db 61 TTACGCCCGCTCTTTTCTTAAGATTATCCCTGAGAAATGCAATATCAAGTCTTATCCA 120
Qy 770 CCAAGGTTGAAATACGCTTGCACAAAGCTGAGCAGGTGACATGACACACCTCGATTATA 829
Db 121 CCAAGTGGTAATACGCTTGCACAAAGCTGAGCAGGTGACATGACACACCTCGATTATA 180
Qy 830 GTGGAAGACCAAGACTGTGTTCCCGAGAGATAGACAGCCAGCTGGAACAGCCCCAGAC 889
Db 181 GTGGAAGACCAAGAGCTGTTCCCGAGAGATAGACAGCCAGCTGGAACAGCCCCAGAC 240
Qy 890 CTTATACCCATCTTCAAGCGGCAAAAGACTGGGATAAATCTGGAAGCTGAAGTCAAAA 949
Db 241 CTTATACCCATCTTCAAGCGGCAAAAGACTGGGATAAATCTGGAAGCTGAAGTCAAAA 300
Qy 950 AGGAGGAGAGAGAGAAACTTGTGCTGATGCTGCTGATGCAATGAAACAAATCTTCGCTGACA 1009
Db 301 AGGAGGAGAGAGAGAAACTTGTGCTGATGCTGCTGATGCAATGAAACAAATCTTCGCTGACA 360
Qy 1010 TCTACAAAGGATGCTGATGAGATATGCGAGGGCCATGATGAAGTCAATTCGTGGAATCAA 1069
Db 361 TCTACAAAGGATGCTGATGAGATATGCGAGGGCCATGATGAAGTCAATTCGTGGAATCAA 420
Qy 1070 ATGCACTGTTCTCAACCAATGGAAGATGTTGGAGCAAAAGAGTGAAGGGAGCC 1129
Db 421 ATGCACTGTTCTCAACCAATGGAAGATGTTGGAGCAAAAGAGTGAAGGGAGCC 480
Qy 1130 CCCTGATGATGAGGAGCTCAAGAAGTGGGAATGCTAAAGTTTGGAC-TGCCCGTCTTTT 1188
Db 481 CTCTGATGATGAGGAGCTTAAGAAGTGGGAATGCTAAAGTTTGGAC-TGCCCGTCTTTT 540
Qy 1189 GTAATCCAGGCTTGGAAACTATGACCTAAATCTGCCCCACCCA--TAGTGCCATGAGCT 1246
Db 541 GTAATCCAGGCTTGGAAATGATGACCTAGTAACGTCACCCATTTGGTGGCCATGAGCT 600
Qy 1247 TGCTTGGTAAAGCTCTGCTTTGTAAGCTTTCTGTATGACACTAGCGTGCTGAG 1306
Db 601 TG---GTTAAGCTTTGTAAGCTTTCTGTATGACAGCTGACGCTGACGCTGCTGAG 657
Qy 1307 TGCTACTGCGCTGAGCTTTTGTAGTGACAT-AGGTATGCTATTATGTGGTATGACCCCGA 1365
Db 658 TGCTACTGCGCTGAGCTTTTGTAGTGACATGCTGATGCTATTCTGTGGTATGACCTGCGA 717
Qy 1366 ACATCTGTTGGTATCAAGACAGCAAACTAATATGATGGAATGCTATCTTTTGGTC 1425
Db 718 AACATTTGGTAAACCAAGACAGCAAACTAATCTGTGGAATGGCTATTCTTTTGGTC 777
Qy 1426 T 1426
Db 778 T 778

RESULT 13
CA215140
LOCUS
DEFINITION SCSBAD1128H12.g Ad1 Saccharum officinarum cDNA clone SCSBAD1128H12
5', mRNA sequence.
ACCESSION CA215140
VERSION CA215140.1 GI:35263122
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 666)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries chat made SUCST
Genet. Mol. Biol..24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 128 row: H column: 12
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .666
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCSBAD1128H12"
/lab_host="DH10B"
/clone_lib="AD1"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucst.lad.ic.unicamp.br/public"

ORIGIN

Query Match 43.4%; Score 631; DB 6; Length 666;
Best Local Similarity 96.5%; Pred. No. 1.3e-152;
Matches 643; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 420 ATGTGATGAGCCGATCGCTGAGGAATCTAGCCAGGACACAGTAAAGAAATGTTGAGGCTAC 479
Db 1 ATGTGATGAGTGCATCGCTGAGGAATCTAGCCAGGACCCAGTAAAGAAAGTTGAGGCTCC 60
Qy 480 TGTGGCTGCTACTATTGAGGACAAAGGAGTTTCAAAATATGGAGAAATACACCCAGCT 539
Db 61 TGTGGCTGCTACTGTGAGGACAAAGGAGTGCACAAATATGGAGAAATACACCCAGCT 120
Qy 540 GATAGAACCCCAAGCAAAACCAAAATACAGGCATGACTACTACACAGTCCACAGAACT 599
Db 121 GGTAGAACCCCAAGCAAAACCTAAATACAGGCATGACTACTACACAGTCCACAGAACT 180
Qy 600 GGTACTGCAATATTTGCTTAAGGGTGTTCCTGCTGATGATGATGATGATGATGATGATG 659
Db 181 GGTGCTGCAATATTTGCTTAAGGGTGTTCCTGCTGATGATGATGATGATGATGATGATG 240
Qy 660 ACAGATGTTGAGTGTATCCATTGAAAGTTCTGGTGAAGAACCATACCATTTTTCAGCCCG 719
Db 241 ACAGATGTTGAGTGTATCCATTGAAAGTTCTGGTGAAGAACCATACCATTTTTCAGCCCG 300
Qy 720 TCTGTTTCTTAAGATTATTCCTTGAGAAATGCAAAATATCAAGTCTTTATCCACCAAGTTGA 779
Db 301 TCTGTTTCTTAAGATTATTCCTTGAGAAATGCAAAATATCAAGTCTTTATCCACCAAGTTGA 360
Qy 780 AATAGCCCTTGCNAAGCTGAGCAGGTGACATGGAACACCTTGGATTTATGTTGGAAGACC 839
Db 361 AATAGCCCTTGCNAAGCTGAGCAGGTGACATGGAACACCTTGGATTTATGTTGGAAGACC 420
Qy 840 AAAGACTGTTCCCAAGAGATAAGCACCCAGCTGAAACAGCCCCCAAGCTTCTATACCC 899
Db 421 AAAGCTGTTCCCAAGAGATAAGCACCCAGCTGAAACAGCCCCCAAGCTTCTATACCC 480
Qy 900 ATCTTCAAGGCGGAAAAAAGACTGGATTAATCTGGAAGCTGAAGTCAAAAAAGGAGGAAA 959
Db 481 ATCTTCAAGGCGGAAAAAAGACTGGATTAATCTGGAAGCTGAAGTCAAAAAAGGAGGAAA 540
Qy 960 GGAAGAAAACTTGTATGTTGATGCTGCAATGCAAACTTCTTCGCTGATCTTACAGGA 1019
Db 541 GGAAGAAAACTTGTATGTTGATGCTGCAATGCAAACTTCTTCGCTGATCTTACAGGA 600

```

QY 1020 TGCTGATGAGATATCGGAGGCCCATGAGTCAATTCGTGAATCAAAATGGCACTGT 1079
Db |||||||
QY 601 TGCTGATGAAGATATCGGNAGGCCATGATGAATCATTCGTGAATCTAAATGGCACTGT 660
Db |||||||
QY 1080 TCTCTC 1085
Db 661 TCTCTC 666

RESULT 14
AW076372 681 bp mRNA linear EST 14-OCT-1999
LOCUS 614066C10.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW076372
VERSION AW076372.1 GI:6031470
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 681)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614066 row: C column: 10.
FEATURES
source
Location/Qualifiers
1..681
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/clone_lib="614 - root cDNA library from Walbot Lab"
/notes="Organ: root; Vector: pBlueScriptII SK+; Site 1:
EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
ORIGIN
Query Match 43.2%; Score 628.8; DB 2; Length 681;
Best Local Similarity 99.7%; Pred. No. 5e-152;
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 735 TATCCCTGAGAATGCAATATATCAAGTCTTATCCACCAAGTTGAATACGCTTGCAAA 794
Db |||||||
QY 19 TATCCCTGAGAATGCAATATATCAAGTCTTATCCACCAAGTTGAATACGCTTGCAAA 78
Db |||||||
QY 795 AGCTGAGCAGGTGACATGGACACCTCGGATTATAGTGGAGACCAAGACTGTTCCCA 854
Db |||||||
QY 79 AGCTGAGCAGGTGACATGGACACCTCGGATTATAGTGGAGACCAAGACTGTTCCCA 138
Db |||||||
QY 855 GAAGATAAGCACGCCAGCTGAACACAGCCCCCAAGACCTTTCATACCCATCTTCAAAGGCGAA 914
Db |||||||
QY 139 GAAGATAAGCACGCCAGCTGAACACAGCCCCCAAGACCTTTCATACCCATCTTCAAAGGCGAA 198
Db |||||||
QY 915 AAAAGACTGGATAAATCGAAGCTGAAGTCAAAAGAGGAGAGAGAAAGAACTTGA 974
Db |||||||
QY 199 AAAAGACTGGATAAATCGAAGCTGAAGTCAAAAGAGGAGAGAGAGAAAGAACTTGA 258
Db |||||||
QY 975 TGCTGATGCTGATTTGACAAATCTTCCGTGACATCTACAAGATGCTGATGAAGATAT 1034
Db |||||||

```

```

Db 259 TGGTGATGCTGCATTGAACAAATTTCTTCGTGACATCTACAAGGATGCTGATGAAGATAT 318
QY 1035 GCGGAGGGCCCATGATGAAGTCAATTCGTGAATCAAAATGSCACTGTTCTCTCAACCAATTC 1094
Db |||||||
QY 319 GCGGAGGGCCCATGATGAAGTCAATTCGTGAATCAAAATGSCACTGTTCTCTCAACCAATTC 378
Db |||||||
QY 1095 GAAAGATGTTGGAGCAAGAAAGGTAGAAGGGAGCCCCCTCGATGGTATGAGGCTCAAGAA 1154
Db 379 GAAAGATGTTGGAGCAAGAAAGGTAGAAGGGAGCCCCCTCGATGGTATGAGGCTCAAGAA 438
QY 1155 GTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTTGGAAATATGA 1214
Db 439 GTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTTGGAAATATGA 498
QY 1215 CCTAATTCGCCACCATAGTCCATGAGCTTGTGTTAAAGTCTCTGCTTTTGTAAAG 1274
Db 499 CCTAATTCGCCACCATAGTCCATGAGCTTGTGTTAAAGTCTCTGCTTTTGTAAAG 558
QY 1275 CTTTCTGATGACAGTCTAGCGTGTCTGAGTGTCTTGCCTAGCTTTTGCATGGACA 1334
Db 559 CTTTCTGATGACAGTCTAGCGTGTCTGAGTGTCTTGCCTAGCTTTTGCATGGACA 618
QY 1335 TAGGTATGCTATTATGTGTTATCGACCCCGAA 1366
Db 619 TAGGTATGCTATTATGTGTTATCGACCCCGAA 650

RESULT 15
AW155786 658 bp mRNA linear EST 04-NOV-1999
LOCUS 614096F11.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW155786
VERSION AW155786.1 GI:6227082
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 658)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614096 row: F column: 11.
FEATURES
source
Location/Qualifiers
1..658
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/clone_lib="614 - root cDNA library from Walbot Lab"
/notes="Organ: root; Vector: pBlueScriptII SK+; Site 1:
EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
ORIGIN
Query Match 43.1%; Score 627; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.5e-151;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 829 AGTGAAGACCAAGACTGTTCCCAAGATAGCAGCCAGCTGAACAGCCCCCAAGA 888

```

Db 9 |||||AGAGACCAAGACTGTTCCCGAGAGATAAGCAGCGCAGCTGAACAGCCCCAAGA 68
QY 889 CCTTCATACCCATCTTCAAAGGCGAAAAAGACTGGGATAAAGCTGGAAGCTGAAGTCAAA 948
Db 69 CCTTCATACCCATCTTCAAAGGCGAAAAAGACTGGGATAAAGCTGGAAGCTGAAGTCAAA 128
QY 949 AAGGAGGAGAGGAAGAAACCTTGATGGTGTGCTGTCATTTGAACAAATTTCTTCGGTGAC 1008
Db 129 AAGGAGGAGAGGAAGAAACCTTGATGGTGTGCTGTCATTTGAACAAATTTCTTCGGTGAC 188
QY 1009 ATCTACAAGGATGCTGATGAAGATATGCGGAGGCGCCATGATGAAGTCATTTCTGGATCA 1068
Db 189 ATCTACAAGGATGCTGATGAAGATATGCGGAGGCGCCATGATGAAGTCATTTCTGGATCA 248
QY 1069 AATGGCACTGTTCTCTCAACCAATTTGAAAGATGTTGGAGCAAAAGAGGTAGAAGGAGC 1128
Db 249 AATGGCACTGTTCTCTCAACCAATTTGAAAGATGTTGGAGCAAAAGAGGTAGAAGGAGC 308
QY 1129 CCCCTGATGGTATGAGGCTCAAGAACTATGACCTAATTTCTGCCACCCATAGTSCCATGAGCTTG 1188
Db 309 CCCCTGATGGTATGAGGCTCAAGAACTATGACCTAATTTCTGCCACCCATAGTSCCATGAGCTTG 368
QY 1189 GTAAATCCAGGCTTGGAACTATGACCTAATTTCTGCCACCCATAGTSCCATGAGCTTG 1248
Db 369 GTAAATCCAGGCTTGGAACTATGACCTAATTTCTGCCACCCATAGTSCCATGAGCTTG 428
QY 1249 CTTGGTTAAGTCTCTGCTTTTGTAAAGCTTCTGTATGACAGTCTAGCGTGTGCTGAGTG 1308
Db 429 CTTGGTTAAGTCTCTGCTTTTGTAAAGCTTCTGTATGACAGTCTAGCGTGTGCTGAGTG 488
QY 1309 CTACTTGCCGTAGCTTTTGATGACATAGGTATGCTATTATGTGGTATCGACCCCGAACC 1368
Db 489 CTACTTGCCGTAGCTTTTGATGACATAGGTATGCTATTATGTGGTATCGACCCCGAACC 548
QY 1369 ATCTGTGGTATCAAGACACCAACTATATGTATGGAATGCCCTATTTCTTTTGGTCTAA 1428
Db 549 ATCTGTGGTATCAAGACACCAACTATATGTATGGAATGCCCTATTTCTTTTGGTCTAA 608
QY 1429 AAAAAAAAAAAAAAAAAAAAAA 1455
Db 609 AAAAAAAAAAAAAAAAAAAAAA 635

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2005, 23:55:26 ; Search time 841.735 Seconds
(without alignments)
10232.696 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcgacgagcgac.....aaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734.4	50.5	1104	12	ADQ37150 Cell prol
2	734.4	50.5	1104	12	ADQ5734 Rice stre
3	571	39.2	571	10	ADQ51379 Maize Rar
4	420	28.9	1366	3	AAC39274 Arabidops
5	419.6	28.8	1324	3	AAC40498 Arabidops
6	418.4	28.8	1361	3	AAC51074 Arabidops
7	417.6	28.7	1077	6	ADG87792 A. thalia
8	393.2	27.0	1117	3	AAC42413 Arabidops
9	388	26.7	388	10	ADF51380 Maize Rar
10	286	19.7	286	6	ABL71197 Corn tass
11	240.2	16.5	581	9	ACL25190 DNA clone
12	221.8	15.2	585	13	ACN47711 Cotton pr
13	217.8	15.0	534	13	ACN59636 Cotton gy
14	216.4	14.9	272	6	ABL75330 Corn tass
15	214.4	14.7	541	9	ACL25189 DNA clone
16	203.6	14.0	571	9	ACL25186 DNA clone
17	201.6	13.9	291	6	ABL73761 Corn tass
18	194.4	13.4	528	13	ACN59570 Cotton gy
19	189.6	13.0	500	9	ACL25185 DNA clone
20	183.8	12.6	573	13	ACN47765 Cotton pr

C	21	175.6	12.1	523	13	ACN61180 Cotton gy
	22	164.4	11.3	272	10	ABX86113 Corn ear-
	23	163	11.2	1535	8	ACD13363 Human DNA
	24	163	11.2	1555	10	ADD29650 Human tum
	25	163	11.2	1776	5	ABV28161 Human pro
	26	163	11.2	1776	5	ABV22337 Human pro
	27	156.4	10.7	1025	3	AAS2451 HTRM clon
	28	156.4	10.7	1185	6	ABL42163 Nucleotid
	29	156.4	10.7	1284	6	ABQ55038 Human ova
	30	156.4	10.7	1366	8	ABZ24700 Human cel
	31	156.4	10.7	2053	2	AAZ27345 Human sec
	32	156.4	10.7	2053	9	ADA07224 Human cDN
	33	156.4	10.7	2053	12	ADN40960 Novel hum
C	34	142.2	9.8	409	8	ABX62506 Arabidops
	35	133.6	9.2	615	13	ADQ55550 Novel can
	36	129.2	8.9	481	10	ADE82158 Arabidops
C	37	111.4	7.7	464	10	ADB50274 Primary r
C	38	111.2	7.6	340	9	ACL25187 DNA clone
C	39	111.2	7.6	412	9	ACL25184 DNA clone
	40	93.2	6.4	672	12	ADJ40154 Plant cDN
	41	90.8	6.2	1242	6	ABZ31779 Candida a
	42	85.4	5.9	431	6	ABL62696 Colon ade
	43	78.2	5.4	518	6	ABQ55648 Human ova
	44	78	5.4	1303	12	ADJ43595 Plant cDN
	45	75.8	5.2	869	4	AAK63551 Human imm

ALIGNMENTS

RESULT 1

ADQ37150

ID ADQ37150 standard; DNA; 1104 BP.

XX AC ADQ37150;

DT 07-OCT-2004 (first entry)

XX Cell proliferation-related nucleic acid sequence #155.

DE cell proliferation related polypeptide; cell proliferation; senescence;

XX differentiation; stress response; ds.

OS Oryza sativa.

XX WO2004061122-A2.

XX PD 22-JUL-2004.

XX PF 23-DEC-2003; 2003WO-US041200.

XX PR 26-DEC-2002; 2002US-0436565P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Cooper B;

XX WPI; 2004-534388/51.

XX New nucleic acid molecule encoding a cell proliferation-related polypeptide, useful for modulating cell proliferation, senescence, differentiation, development, and stress response in plants, and for producing enhanced food crops.

XX Claim 57; SEQ ID NO 309; 408pp; English.

XX The present invention relates to an isolated nucleic acid molecule encoding a cell proliferation-related polypeptide. The nucleic acid molecule and the encoded polypeptide, and methods are useful for modulating cell proliferation, senescence, differentiation, development, and stress response in plants, and for producing enhanced food crops. The present sequence represents a cell proliferation-related nucleic acid sequence. The present sequence is published separately from the main body

CC of the specification as EPO data.

XX Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

SQ Query Match 50.5%; Score 734.4; DB 12; Length 1104;
Best Local Similarity 80.5%; Pred. No. 3.3e-151;
Matches 893; Conservative 0; Mismatches 196; Indels 21; Gaps 2;

Qy 73 AGGGCGCCATGGCGCGTGGATCTGGAGAGCAAGCGGCGCTTCGTGACGAC 132
Db 1 ATGGCAACCGCGCGTGGATCTGGAGAGCAAGCGGCGCTTCGTGACGAC 60

Qy 133 GACTTCGAGTGGCCACCGAGCTTCAAGCAGGCGCATGACGCGCGCGCGCCGCC 192
Db 61 GACTTCGAGTGGCGCGCGAGCTTCAAGCAGGCGCATGACGCGCGCGCGCCGCC 120

Qy 193 GACTTCATGCGGAGCGCGCGAGGCGCATGACGCGCGCGCGCGCGCGCGCGCG 252
Db 121 GACTTCATGCGGAGCGCGCGAGGCGCATGACGCGCGCGCGCGCGCGCGCGCG 180

Qy 253 GCGGATGCTAACAAAGCAATGAGCTTGCCTATGATGATGATGATGATGATGAT 312
Db 181 GCTGATGCTAACAAAGCAATGAGCTTGCCTATGATGATGATGATGATGATGAT 240

Qy 313 GGTGCTGATGCTAACAAAGCAATGAGCTTGCCTATGATGATGATGATGATGAT 372
Db 241 GGTGCTGATGCTAACAAAGCAATGAGCTTGCCTATGATGATGATGATGATGAT 300

Qy 373 TCTTCTTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 432
Db 301 TACTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 360

Qy 433 ATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 477
Db 361 ATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 420

Qy 478 ACTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 537
Db 421 TCTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480

Qy 538 GTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 597
Db 481 ATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 534

Qy 598 GTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 657
Db 535 GTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 594

Qy 658 GAACAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 717
Db 595 GAACAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 654

Qy 718 CGTCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 777
Db 655 CGTCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 714

Qy 778 GAAATAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 837
Db 715 GAAATAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 774

Qy 838 CCAAGAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 897
Db 775 CCAAGAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 834

Qy 898 CCATCTTCAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
Db 835 CCATCTTCAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 894

Qy 958 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1017
Db 895 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 954

Qy 1018 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077

Db 955 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014

Qy 1078 GTTCTCTCAACCAATTGGAAGATGTTGGAGCAAGAGCTAGAAGGAGGAGGAG 1137
Db 1015 GTTCTCTCAACCAATTGGAAGATGTTGGAGCAAGAGCTAGAAGGAGGAGGAG 1074

Qy 1138 GGTATGAGCTCAAGAGTGGGAATCTAA 1167
Db 1075 GGTATGAGCTCAAGAGTGGGAATCTAA 1104

RESULT 2
ADQ15734
ID ADQ15734 standard; DNA; 1104 BP.
XX AC ADQ15734;
XX DT 07-OCT-2004 (first entry)
XX DE Rice stress-related protein coding sequence #72.
XX KW rice; stress-related protein; plant maturation; plant development;
XX KW plant proliferation; plant senescence; plant disease-resistance;
XX KW plant stress response; transgenic plant; pest tolerance;
XX KW herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
XX KW improved nutritional value; increased yield; increased proliferation;
XX KW gene; ds.
XX OS Oryza sativa.
XX PN WO2004061080-A2.
XX PD 22-JUL-2004.
XX PF 23-DEC-2003; 2003WO-US041098.
XX PR 26-DEC-2002; 2002US-0436564P.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Cooper B;
XX WPI; 2004-534374/51.
XX P-PSDB; ADQ15735.
XX PT New isolated nucleic acids and proteins, useful for producing transgenic
XX PT plants having improved properties, e.g. tolerance to pests, herbicides,
XX PT or biotic or abiotic stresses, improved nutritional value, or increased
XX PT yield or proliferation.
XX PS Claim 45; SEQ ID NO 143; 551pp; English.
XX CC The invention comprises the amino acid and coding sequences of rice
XX CC stress-related proteins. The DNA and protein sequences of the invention
XX CC are useful for regulating and controlling plant maturation and
XX CC development, including proliferation, senescence, disease-resistance, or
XX CC stress response. They are also useful for producing transgenic plants
XX CC having improved properties, e.g. tolerance to pests, herbicides, or
XX CC biotic or abiotic stresses, improved nutritional value, increased yield
XX CC or proliferation, or improved structure causing less loss from lodging or
XX CC shattering. The present DNA sequence encodes a rice stress-related
XX CC protein of the invention.
XX SQ Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

Query Match 50.5%; Score 734.4; DB 12; Length 1104;
Best Local Similarity 80.5%; Pred. No. 3.3e-151;
Matches 893; Conservative 0; Mismatches 196; Indels 21; Gaps 2;

Qy 73 AGGGCGCCATGGCGCGTGGATCTGGAGAGCAAGCGGCGCTTCGTGACGAC 132
Db 1 ATGGCAACCGCGCGTGGATCTGGAGAGCAAGCGGCGCTTCGTGACGAC 60


```
Db 181 |CCCCCGCCCGCCGACCTCTATGCGGACCGCGCCGCGCACATCAAGCTCGGCAACTAC 240
Qy 241 |ACTGAGGCTGTGGCGGATGCTAAACAAGCAATTGAGCTTGTATCTATGATCATATAAGCT 300
Db 241 |ACTGAGGCTGTGGCGGATGCTAAACAAGCAATTGAGCTTGTATCTATGATCATATAAGCT 300
Qy 301 |TACTACCGGAAGGTGTGTCATGATTAAGCTTGAAGAATACCAAACTGCAAAAGGCTGCT 360
Db 301 |TACTACCGGAAGGTGTGTCATGATTAAGCTTGAAGAATACCAAACTGCAAAAGGCTGCT 360
Qy 361 |CTTGAGTGGGTTCTTCTATGCAATCAGGCAATCAAGTTTGTCTCTATTGAAAGGAA 420
Db 361 |CTTGAGTGGGTTCTTCTATGCAATCAGGCAATCAAGTTTGTCTCTATTGAAAGGAA 420
Qy 421 |TGTGATGAGGCATCGCTGAGGAATCTAGCCAGGCACCAAGTAAAGAATGTTGAGGCTACT 480
Db 421 |TGTGATGAGGCATCGCTGAGGAATCTAGCCAGGCACCAAGTAAAGAATGTTGAGGCTACT 480
Qy 481 |GTGCTGCTACTATTGAGGACCAAGGAGATTTCACAAATATGGAGAATACACCAACCAAGTG 540
Db 481 |GTGCTGCTACTATTGAGGACCAAGGAGATTTCACAAATATGGAGAATACACCAACCAAGTG 540
Qy 541 |ATAGAACCCCGCCAGCAACCAACCAATACAGGC 571
Db 541 |ATAGAACCCCGCCAGCAACCAACCAATACAGGC 571

RESULT 4
AAC39274
ID AAC39274 standard; DNA; 1366 BP.
XX
AC AAC39274;
XX
AT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 24023.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 21-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
```

PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	<div>Query Match 28.9%; Score 420; DB 3; Length 1366; Best Local Similarity 63.0%; Pred. No. 3.7e-82; Matches 683; Conservative 0; Mismatches 395; Indels 6; Gaps 2;</div>		
PR	05-AUG-1999;	99US-0147192P.	Qy	88	CGCTCGGATCTGGAGAGCAAGGCCCTTCGTCGACGAGCTTCGAGCTGGCC 147
PR	05-AUG-1999;	99US-0147260P.	Db	60	GCGAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTCTAGATGATGACTTCGATCTTGT 119
PR	06-AUG-1999;	99US-0147303P.	Qy	148	ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCGCCACGCCGACCTCTATGCCGAC 207
PR	06-AUG-1999;	99US-0147416P.	Db	120	GTTGACTTTATCTCCAAAGCCATTGACTTGGATCCCAATTGCGCGCCCTTCTTCGCCGAT 179
PR	09-AUG-1999;	99US-0147493P.	Qy	208	CGCGCCAGCGCACATCAAGCTCGCAACTACACTGAGGCTGTGGCGGATGCTAAACAAA 267
PR	10-AUG-1999;	99US-0148171P.	Db	180	CGTGCTCAGGCCAATCAAAATCGATTAACCTTCTGTAAGCTTTGTAGATGCGAACA 239
PR	11-AUG-1999;	99US-0148319P.	Qy	268	GCAATTGAGCTTGATCTCTATGATGATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
PR	12-AUG-1999;	99US-0148341P.	Db	240	GCCATTGAGTTGGAGCCAAACCTTGGCAAAAGCCCTATCTCAGAAAGGCGACTGTTGTATG 299
PR	13-AUG-1999;	99US-0148565P.	Qy	328	AAGCTTGAAGAATACCAAACTGCAAAAGCTGCTCTTGTAGTTGGTTCCTTTATGCATCA 387
PR	13-AUG-1999;	99US-0148684P.	Db	300	AAGCTAGAAGAATATAGTACTGTCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGACCG 359
PR	16-AUG-1999;	99US-0149336P.	Qy	388	GGCGATTCAAGGTTGCTCGTCTATTGAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
PR	17-AUG-1999;	99US-0149175P.	Db	360	AATGAACCAAAAGTTTAAAGAAGATGATAGTGAATGCGATCTTCGTATTGCAAGAAGAG 419
PR	18-AUG-1999;	99US-0149426P.	Qy	448	AGCCAGGCACCAAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTATGAGGACAGGAG 507
PR	20-AUG-1999;	99US-0149722P.	Db	420	A---AAGATTGTTGAGCCGATGCCACCGAGTTGCTTCAAGCTCTACAACACCCTA 476
PR	20-AUG-1999;	99US-0149723P.	Qy	508	GATTTCAAAATATGGAGAATACACCACAGTGATAGAACCCCAAGCAAAACCAAAATAC 567
PR	20-AUG-1999;	99US-0149922P.	Db	477	GCAACGGGAAGCTGATGCTCTCTCTGTTCCAAATTCCTGCAGCACCTGCGCAAAACCGATGTC 536
PR	23-AUG-1999;	99US-0149902P.	Qy	568	AGGCATGACTACTACACAGTGCACAGAGTGCTACTGACAAATATTTTCTTAAGGGTGT 627
PR	23-AUG-1999;	99US-0150566P.	Db	537	AGGCACGAGTTCTACCAAGAAACCAAGAAAGCGGTGGTGACAAATTTTCGCCAAAAAGTA 596
PR	26-AUG-1999;	99US-0150884P.	Qy	628	CCTGCTGATAGTAGTCAATTGATTTTGGTGAACAGAGTGTGAGTGATCCATTCAAGT 687
PR	27-AUG-1999;	99US-0151066P.	Db	597	CCTAAGGAGAAGCTTACTGTCGAGTTTGGTGAGCAGATCTTGAAGTTGTTGTTGATGTT 656
PR	27-AUG-1999;	99US-0151080P.	Qy	688	CCTGCTGAAGAACCATACCATTTTTCAGCCCGTCTGTTTTCTTAAGATTATCCCTGAGAAA 747
PR	30-AUG-1999;	99US-0151303P.	Db	657	GCTGGAGAGGAAGCTTATCATCTCCAGCGAGATTGTTTCGGGAAGATATATACCAGAGA 716
PR	31-AUG-1999;	99US-0151338P.	Qy	748	TGCAATATCAAGTCTTATCCACCAAGTTGAAATACGCTTGCCTGCAAAAGCTGAGCAGTG 807
PR	01-SEP-1999;	99US-0151438P.	Db	717	TGCAGATTGGAATGTTGTCGACCAAAATGTCGATCCGTTCTTGGCAAGAGAGATATATC 776
PR	07-SEP-1999;	99US-0151930P.	Qy	808	ACATGGACAACCTGGATTATAGTGAAGACCAAGAGCTGTTTCCCAAGATAGACGACG 867
PR	10-SEP-1999;	99US-0153070P.	Db	777	ACCTGGGCTCCCTTGATATGTTAAAGGGCAAGAGTGTGTTTCCCAAAACCCATCTCTCA 836
PR	13-SEP-1999;	99US-0153759P.	Qy	868	CCAGCTGAAGACAGCCCAAGACCTTTCATACCCATCTTCAAAAGGGCAAAAGACTGGGAT 927
PR	15-SEP-1999;	99US-0154018P.	Db	837	TCAGC---GCTGTGCGAGAGACCAAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 893
PR	16-SEP-1999;	99US-0154033P.	Qy	928	AACTGGAGCTGAAGTCAAAAAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 987


```
Db      1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCCACAACTGGAAGAAGTTGGG 1073
Qy      1108 GCAAGAAGGTAGAGGAGCCCTCATGCTATGGAGCTCAAGAGTGGGAATACTAA 1167
Db      1074 ACTAAGAAGTGGAGAGCACTCCACCATGGATGGAGCTCAAGAAGTGGGAGTATTGA 1133
Qy      1168 AGTT 1171
Db      1134 TCCT 1137

RESULT 6
AAC51074
ID AAC51074 standard; DNA; 1361 BP.
AC AAC51074;
XX
XX
DT 18-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67179.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144322P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 23-JUL-1999; 99US-0145224P.
XX 26-JUL-1999; 99US-0145276P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 28-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146389P.
XX 03-AUG-1999; 99US-0147038P.
XX 04-AUG-1999; 99US-0147204P.
XX 04-AUG-1999; 99US-0147302P.
XX 05-AUG-1999; 99US-0147192P.
XX 06-AUG-1999; 99US-0147260P.
XX 06-AUG-1999; 99US-0147303P.
XX 06-AUG-1999; 99US-0147416P.
```



```
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 19-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 27.0%; Score 393.2; DB 3; Length 1117;
Best Local Similarity 62.7%; Pred. No. 2.6e-76;
Matches 677; Conservative 0; Mismatches 373; Indels 30; Gaps 3;

QY 88 CGCTCGGATCTGGAGAGCAGCCAGAGGCTTCGTGAGCAGCAGCTTCGAGCTGGCC 147
DB 68 GCGAGGAGCTTGCTGATGAAGCTAAAGAGGCTTCGTAGAGCAGTCTCGATGTTGCT 127
QY 148 ACCGAGCTCTACAGCAGGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 207
DB 128 GTTGACTGTACTCCAAAGCCATTGACTTGGATCTTAATGCGCTGAGTCTTCGCTGAT 187
```

```
QY 208 CGCGCCAGGCGCACATCAAGCTCGCAACATACTAGAGGCTGTGGCGATGCTAACAA 267
DB 188 CGTGTCTAGGCTATATCAAACTCGAAAGCTTCACTGAGGCGTGCAGATGCGCAACA 247
QY 268 GCAATTGAGCTTGATCCTATGATGATGATGATGATGATGATGATGATGATGATGAT 327
DB 248 GCAATTGAGCTTGATCCTTCAATTGACCAAGGCTTACCTTAAGAAAAGGAACCTG 307
QY 328 AAGCTTTGAAGAATACCAAACTGCAAGGCTGCTCTTGAAGTTGGGTCTTCTTTAT 387
DB 308 AAGCTTTGAAGATATCGGACTGCTTAAACAGCTCTTGAAGGGTGGCTCTATACG 367
QY 388 GCGGATTCAGGTTTGTCTGCTTATTTGAAGGAATGATGAGCGCATCGCTGAGGA 447
DB 368 AGTGAATCCAAATTTTGAAGATTGATAGATGAATGCAATTTTCTAATCACAGA 424
QY 448 AGCCAGGACACAGTAAGAATTTGAGGCTACTGTGGCTGCTATCTATTGAGGACA 507
DB 425 -----GAGAAAAGATTTGGTTCAACCCGGTGCCTTCGACTTTGCTTCAAGT 471
QY 508 GATTTTCACAATATGAGAGATACACCAAGCTGATAGAACCCCAAGCAAAACCAAA 567
DB 472 GACAGCACA-----CCAGTATCTGAATTTGATGTACCCCTACAGCCAAGTAC 520
QY 568 AGGCATGACTACTACACAGTGCACAGAGTGGTACTGACAAATTTTGTAAAGGTT 627
DB 521 AGGCACGAGTACTACCAAAAGCAGAGGAAGTTGTGGTAACTGTTTTTGCAAAAG 580
QY 628 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
DB 581 CCAAGCAGAAATGTTAAACATGACTTTTGGTGAACAAATTTCTGAGTGTGTTGAT 640
QY 688 CTGTGTGAAGAACCATACATTTTTCAGCCCGTCTGTTTTCTTAAGATTATCCTCAG 747
DB 641 CTTGAGAGGATGCTGATTTATCTCAACCGAGATGTTTGAAGATATATACCAGAT 700
QY 748 TGCAAAATATCAAGTCTTATCCCAAGGTTGAAATACGCTTGCAAAAGCTGAGCAG 807
DB 701 TGCAAAATATGAAGTATTGTCGACCAAAATTCAGATCTGCTTGTGAAAAGCCGAT 760
QY 808 ACATGGACACCCCTGGATTTAGTGAAGACCAAGACTGTTCCCAAGATAGCAGC 867
DB 761 ACATGGGCTCCTCGAACACGCGCAAGGGCCGCGTTTTGCGCAAGCCCTAAATG 820
QY 868 CCAGCTGAAACAGCCCAAGACCTTTCATACCCATCTTCAAAGGCGAAAAGAGCTGG 927
DB 821 TCAGAGTTTCA---CAGAGGCCAGCTTATCCATCTCTTAAGAAAGTTAAGGACTGG 877
QY 928 AAATGGAAGCTGAAGTCAAAAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 987
DB 878 AAGCTGGAAGCTGAAGTGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 937
QY 988 TTGAACAAATTTCTTCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1047
DB 938 TTGAACAAATTTCTTCGCGGAGATATACCAAGATCGGATGAGGATATGAGGCG 997
QY 1048 ATGAAGTCAATTCGTGGATCAAAATGGCACTGTTCTTCAACCAATTTGGAAGAT 1107
DB 998 AGCAATCATTTTGGATCGAATGGGACAGTGTCTCAACAAACTGCAAGAGGTTGG 1057
QY 1108 GCAAGAAGGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1167
DB 1058 ACTAAGACAAATCGAGAGTACTCTCTCGGATGGCATGGAGCTCAAGAAATGGAG 1117

RESULT 9
ID ADF51380
XX ADF51380 standard; cDNA; 388 BP.
AC ADF51380;
XX
DT 12-FEB-2004 (first entry)
XX
```

DE Maize Rar interactor 2 (RarInt2) DNA seq id 3.
XX plant protectant; gene therapy; Rar1; disease resistance; plant;
KW immunogen; maize; Rar interactor 2; RarInt2; gene; ss.
XX
OS Zea mays.
XX
PN US2003167504-A1.
XX
PD 04-SEP-2003.
XX
PF 12-FEB-2002; 2002US-00074473.
XX
PR 12-FEB-2001; 2001US-0268157P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Crane EH;
XX
XX WPI; 2003-898130/82.
XX
XX New maize Rar1-interactor nucleic acid useful for enhancing disease
PT resistance in plants, as a probe or amplification primer, in recombinant
PT expression of polypeptides, or as an immunogen in the preparation or
PT screening of antibodies.
XX
XX Claim 1; SEQ ID NO 3; 30pp; English.
XX
XX The invention describes an isolated nucleic acid capable of interacting
CC with Rar1. The nucleic acid comprises a member selected from: (a) a
CC polynucleotide comprising a sequence of 388 (S1), 571 (S2) or 691 (S3) bp
CC fully defined in the specification; (b) a polynucleotide amplified from a
CC Zea mays nucleic acid library using primers that selectively hybridise,
CC under high stringency conditions, to loci within the polynucleotide in
CC (a); (c) a polynucleotide encoding a maize Rar1-interactor protein; (d) a
CC polynucleotide having at least 80 or 90% sequence identity to (a); (e) a
CC polynucleotide which hybridises under high stringency conditions to (a);
CC and (f) a polynucleotide complementary to any of the above
CC polynucleotides. The nucleic acid molecule and methods are useful in
CC enhancing disease resistance in plants. The nucleic acid molecule may
CC also be used as a probe or an amplification primer, in recombinant
CC expression of polypeptides, or as an immunogen in the preparation and/or
CC screening of antibodies. This sequence represents a maize Rar interactor
CC 2 (RarInt2) polynucleotide.
XX
SQ Sequence 388 BP; 102 A; 76 C; 93 G; 117 T; 0 U; 0 Other;

Query Match 26.7%; Score 388; DB 10; Length 388;
Best Local Similarity 100.0%; Pred. No. 2.7e-75; Mismatches 0; Indels 0; Gaps 0;
Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1043 CCATGATGAAGTCATTGCGTGAATCAATGGCACTGTTCTCTCAACCAATTGGAAGATG 1102
DB 1 CCATGATGAAGTCATTGCGTGAATCAATGGCACTGTTCTCTCAACCAATTGGAAGATG 60

QY 1103 TTGGAGCAAGAGGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1162
DB 61 TTGGAGCAAGAGGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 120

QY 1163 ACTAAGTTTGGAGTCCCGCTTTTGTAAATCCAGGCTCTTGGAACTATGACCTAATTC 1222
DB 121 ACTAAGTTTGGAGTCCCGCTTTTGTAAATCCAGGCTCTTGGAACTATGACCTAATTC 180

QY 1223 TGCCCAACCATAGTCCATGAGCTTGCTTGGTAAAGTCTCTGCTTTTGAAGCTTTCTGT 1282
DB 181 TGCCCAACCATAGTCCATGAGCTTGCTTGGTAAAGTCTCTGCTTTTGAAGCTTTCTGT 240

QY 1283 ATGACAGTCTAGCGTGTGTCTGAGTGTCTACTCTTCCGCTAGCTTTTGTAGACATAGGTATG 1342
DB 241 ATGACAGTCTAGCGTGTGTCTGAGTGTCTACTCTTCCGCTAGCTTTTGTAGACATAGGTATG 300

QY 1343 CTATTATGTGTATCGACCCCAACCTCTGTGTATCAAGACAGCAACTATAATGTA 1402
DB 1343 CTATTATGTGTATCGACCCCAACCTCTGTGTATCAAGACAGCAACTATAATGTA 1402

DB 301 CTATTATGTGTATCGACCCCAACCTCTGTGTATCAAGACAGCAACTATAATGTA 360
QY 1403 TGGATGCTATTTCTTTTGGTCTTAAAA 1430
DB 361 TGGATGCTATTTCTTTTGGTCTTAAAA 388

RESULT 10
ABL71197
ID ABL71197 standard; cDNA; 287 BP.
XX
AC ABL71197;
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:571.
XX
KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS Zea mays.
XX
PN US2001051335-A1.
XX
PD 13-DEC-2001.
XX
PF 16-APR-1999; 99US-00294093.
XX
PR 21-APR-1998; 98US-0082567P.
XX
PA (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
XX WPI; 2002-163647/21.
XX
XX Novel purified corn tassel-derived polynucleotide useful for determining
PT altered gene expression, to recover regulatory elements and to follow
PT inheritance of desirable characteristics through hybrid breeding
PT programs.
XX
XX Claim 1; SEQ ID NO 571; 201pp; English.
XX
XX The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful in
CC the evaluation, and alteration of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality and yield, and as molecular markers for studying inheritance of
CC multigene traits in a plant breeding program. (I) can be used to produce
CC a tassel-specific profile of gene transcription, a transcript image, to
CC clone regulatory elements for use in transformation vectors, to express a
CC polypeptide, to identify, isolate or extend identical or related corn
CC tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences
XX
SQ Sequence 287 BP; 89 A; 60 C; 70 G; 68 T; 0 U; 0 Other;

Query Match 19.7%; Score 286; DB 6; Length 287;
Best Local Similarity 100.0%; Pred. No. 6e-53; Mismatches 0; Indels 0; Gaps 0;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 ACCTTACTACCGAAAGGTGCTGCATGCAATTAAAGCTTGAAGAAATACCAAACTGCAAGGC 356
Dd 1 AGCTTACTACCGAAAGGTGCTGCATGCAATTAAAGCTTGAAGAAATACCAAACTGCAAGGC 60
QY 357 TGCCTCTGAGTTGGTCTCTTCTATGTCATCAGCGGATCAAGTTTCTGCTCTATTGAA 416
Dd 61 TGCCTCTGAGTTGGTCTCTTCTATGTCATCAGCGGATCAAGTTTCTGCTCTATTGAA 120
QY 417 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACAGTAAGAAATGTTGAGGC 476
Dd 121 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACAGTAAGAAATGTTGAGGC 180
QY 477 TACTGTGGCTGCTACTATTGAGGACAAAGGAGGATTTCAAAAATATGGAGAATACACCACC 536
Dd 181 TACTGTGGCTGCTACTATTGAGGACAAAGGAGGATTTCAAAAATATGGAGAATACACCACC 240
QY 537 ACTGATAGAACCCCAAGCAAAACCAAAATACAGGATGACTACTAC 582
Dd 241 ACTGATAGAACCCCAAGCAAAACCAAAATACAGGATGACTACTAC 286

RESULT 11

ACL25190/c
ID ACL25190 standard; DNA; 581 BP.

XX AC
XX ACL25190;

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DE

XX DNA clone originating in barley containing SNP encoding sequence #15181.

XX KW

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis; gene; ss.

XX OS

XX Hordeum vulgare; var. (cul.Haruna Nijo).

XX PN

XX WO2003057877-A1.

XX PD

XX 17-JUL-2003.

XX PF

XX 16-DEC-2002; 2002WO-IB005403.

XX PR

XX 20-DEC-2001; 2001JP-00387059.

XX PR

XX 20-DEC-2001; 2001JP-00387131.

XX PR

XX 20-DEC-2001; 2001JP-00403299.

XX PR

XX 20-DEC-2001; 2001JP-00403300.

XX PR

XX 27-SEP-2002; 2002JP-00327515.

XX PA

XX (UYN1-) UNIV JAPAN OKAYAMA.

XX PI

XX Sato K, Takeda K, Kohara Y;

XX XX

XX WPI; 2003-587127/55.

XX DR

XX Single nucleotide polymorphism sites in barley varieties and DNA sequences containing them for analysis and identification of barley varieties and production of barley transformants with desired characteristics.

XX XX

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX PS

XX The present invention relates to oligonucleotide clones originating in barley (Hordeum vulgare) which contain single nucleotide polymorphisms (SNP). The oligonucleotides may be used for analysis of SNPs among barley varieties, identification of particular varieties and genotype-phenotype analysis, isolation of specific genes and creation of new varieties by transformation of barley varieties with them and production of new barley varieties with desired properties. The present sequence represents an oligonucleotide clone DNA sequence featured in the specification. The sequence data for this patent did not form part of the printed form specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to standardise OS field)

XX SQ Sequence 581 BP; 157 A; 151 C; 101 G; 172 T; 0 U; 0 Other;

Query Match 16.5%; Score 240.2; DB 9; Length 581;
Best Local Similarity 79.6%; Pred. No. 8.4e-43;
Matches 284; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 847 GTTCCCAAGAGATAAGCAACGCCAGCTGAACAGCCCAAGACCTTCATACCCATCTTCA 906

Dd 569 GCTCTCTCAGAGATAAATGTACCACTGAATCAGCCCAAGAGCCATCTTATCTCTTCA 510

QY 907 AAGCGCAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAAGGAAGAA 966

Dd 509 AAATCCAAAAGGACTGGGATAAAGCTTGAGGCTGAAGTGAATAACAGAGGAGATGAG 450

QY 967 AAACCTTGATGTGATGCTGCATTTGAACAAATCTTCCTGTGACATCTCAAGATGCTGAT 1026

Dd 449 AAACCTTGACGGTGTGCTGCATTTGAACAAATCTTCCTGTGACATCTCAAGATGCTGAT 390

QY 1027 GAAGATATGCGGAGGCCCATGATGAAGTCATTCTGTGAATCAAAATGGCAGCTTCTCTCA 1086

Dd 389 GAAGATATGCGTAGAGCAATGATGAAGTCATTCTGTGAATCAAAATGGCAGCTTCTCTCA 330

QY 1087 ACCAATTGGAAGATGTTGGAGCAAGAGGTAGAAGGGAGGCCCTCTGATGGTATGGAG 1146

Dd 329 ACCAATTGGAAGATGTTGGAGCAAGAGGTAGAAGGGAGGCCCTCTGATGGTATGGAG 270

QY 1147 CTCAAGAGTGGGAATCTAAAGTTTGGAGTGCCTGCTTTTGAATCCAGGCTTT 1203

Dd 269 CTCAAGAGTGGGAATCTAAAGTTTGGAGTGCCTGCTTTTGAATCCAGGCTTT 213

RESULT 12

ACN47711

ID ACN47711 standard; cDNA; 585 BP.

XX AC

XX ACN47711;

XX DT

XX 02-DEC-2004 (first entry)

XX XX

XX Cotton primed seed EST Clone ID: LI3825-014-Q1-K6-D9, SEQ:2492.

XX KW

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; seed; variety DP50B; library LI3825; molecular tag; molecular marker; genetic mapping; molecular mapping; seed germination; plant growth; plant quality; plant yield; plant breeding; tissue printing; ss.

XX OS

XX Gossypium hirsutum.

XX PN

XX US2004123340-A1.

XX XX

XX 24-JUN-2004.

XX PF

XX 12-DEC-2001; 2001US-00021323.

XX PR

XX 14-DEC-2000; 2000US-0255619P.

XX XX

XX (DEIK/) DEIKMAN J.

XX (FENG/) FENG P C C.

XX (FINC/) FINCHER K L.

XX (ZIEG/) ZIEGLER T E.

XX XX

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX PI

XX WPI; 2004-479808/45.

XX DR

XX New isolated nucleic acid molecule that encodes a plant protein or its fragment, useful for isolating a variety of agronomically significant genes associated with plant growth, quality or yield, and as molecular tags to map genes.

XX PT

XX XX

PS Claim 1; SEQ ID NO 2492; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs);

CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated

CC from primed or non-primed seeds from variety DP50B, mature seeds from

CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeceum

CC tissue, developing fibres, carpel walls and septa from variety

CC Nucotton33B. The invention also relates to substantially purified

CC proteins or their fragments encoded by nucleic acid molecules of the

CC invention, and to transformed plants having a nucleic acid construct

CC comprising a nucleic acid of the invention. The cotton ESTs are useful as

CC molecular tags to isolate genetic regions, to isolate genes, to map

CC genes, to determine gene function and to determine whether genes are

CC members of a particular gene family. The nucleic acid molecules may be

CC used for isolating a variety of agronomically significant genes

CC associated with plant growth, quality, yield, and could also serve as

CC links in metabolic and catabolic pathways. The nucleic acid molecules are

CC also useful for identifying genes important in initiating and maintaining

CC seed germination or that may be used to mitigate stresses encountered

CC during seed germination. The ESTs additionally enable the acquisition of

CC promoters and cis-regulatory elements which will be useful to express

CC agronomically significant genes in these tissues and/or other tissues,

CC and also permits the acquisition of molecular markers useful in breeding

CC schemes, genetic and molecular mapping, and in cloning of agronomically

CC significant genes. The nucleic acid molecules are further useful for

CC detecting the expression level or pattern of a protein or mRNA and for

CC detecting the presence or quantity of a protein by tissue printing. The

CC present sequence represents a specifically claimed EST isolated from a

CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence

CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from the US patent office at

CC seqdata.uspto.gov/sequence.html?DocID=US20040123340

XX

Seq Sequence 585 BP; 186 A; 117 C; 134 G; 148 T; 0 U; 0 Other;

Query Match 15.2%; Score 221.8; DB 13; Length 585;

Best Local Similarity 62.7%; Pred. No. 9.2e-39;

Matches 362; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

QY 229 CTCGGCAACTACTGAGGCTGTGGCGATGCTAAACAAAGCAATGAGCTTGATCCTATG 288

DB 1 CTCACAAATCTCCCTGACGCTGTGGCAGAGCTAACAAAGCAATGAGTTGGATCCGTC 60

QY 289 ATGCATTAAGCTTACTACCGGAAGGTGCTGCATGCATTAAGCTTGAAGATACCAACT 348

DB 61 ATGCTAAATCCTACTTGGTAAAGCTACTGCGTGTATGAAGCTTGGAGGATCAAACT 120

QY 349 GCAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGTCATCAGGCGATTCAAGGTTTCTCGT 408

DB 121 GCTAAGGCTGGTTGGAGACTGGGGCTGCTTTGGCACCAGAGACTCAGATTTTCCAAG 180

QY 409 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCCACCAAGTAAGAT 468

DB 181 TTGATTAAAGAAATGTGAAGAGCGCATTCAGAGGAATGTTGAGTTACCAAGGAGAGC 240

QY 469 GTTGAGGCTACTGTGGCTGCTACTATTGAGGACAGAGGAGATTTCACAAATATGAGAT 528

DB 241 TTGGAAGAGGTGCCAACAAATTTGTACTCGGAAAGACATCTCTTCTGTCCAGGATATA 300

QY 529 ACACCACAGTGA---TAGAACCCCAAGCAACCAAAATACAGGCATGACTACTACAC 585

DB 301 CCGATCCCATGACTGTTGGAGCACCACCAATCACTTACAGGCATGAATTTTATCAG 360

QY 586 AGTGCCACAGAGTGGTACTGACAAATATTTGTAAGGGTGTTCCTGCTGATGTAGTCTC 645

DB 361 AAACACAGAGGAAGTGGTTGTACATAATTTGCAAGGGAATACACCGGAGTGTGTAA 420

QY 646 ATTGATTTTGGTGAACAGATGTTAGTGTATCCATTGAAGTTCTCGGTGAAGAACCATAC 705

DB 421 GTTATTATGTTGGTGAACAAATACTAAGTGTGTGTCAATGCACCCGGCAAGATCTTAT 480

QY 706 CATTTTCAGCCCGTCTCTTTCTAAGATTATCTCCTGAGAAATGCAATATCAAGTCTTA 765

DB 481 CATCTCCAACCTCGCTTATTTTGAAGAATAATACCTGACAAGTCAGATATGATGTTTG 540

QY 766 TCCACCAAGGTTCAATATAGCTTCCAAAAGCTGACC 802

DB 541 TCAACCAAAATTTGAATTTAGGTGTAGCAAAAGCTGAAC 577

RESULT 13

ACN59636

ID ACN59636 standard; cDNA; 534 BP.

XX ACN59636;

XX 02-DEC-2004 (first entry)

XX Cotton gynoeceum tissue EST Clone ID: LIB3829-033-Q1-K6-G10, SEQ:14417.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoeceum;

XX variety Nucotton33B; library LIB3829; molecular tag; molecular marker;

XX genetic mapping; molecular mapping; seed germination; plant growth;

XX plant quality; plant yield; plant breeding; tissue printing; ss.

XX Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

XX (FENG/) FENG P C C.

XX (FINC/) FINCHER K L.

XX (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its

XX fragment, useful for isolating a variety of agronomically significant

XX genes associated with plant growth, quality or yield, and as molecular

XX tags to map genes.

XX Claim 1; SEQ ID NO 14417; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs);

XX ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated

XX from primed or non-primed seeds from variety DP50B, mature seeds from

XX variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeceum

XX tissue, developing fibres, carpel walls and septa from variety

XX Nucotton33B. The invention also relates to substantially purified

XX proteins or their fragments encoded by nucleic acid molecules of the

XX invention, and to transformed plants having a nucleic acid construct

XX comprising a nucleic acid of the invention. The cotton ESTs are useful as

XX molecular tags to isolate genetic regions, to isolate genes, to map

XX genes, to determine gene function and to determine whether genes are

XX members of a particular gene family. The nucleic acid molecules may be

XX used for isolating a variety of agronomically significant genes

XX associated with plant growth, quality, yield, and could also serve as

XX links in metabolic and catabolic pathways. The nucleic acid molecules are

XX also useful for identifying genes important in initiating and maintaining

XX seed germination or that may be used to mitigate stresses encountered

XX during seed germination. The ESTs additionally enable the acquisition of

XX promoters and cis-regulatory elements which will be useful to express

XX agronomically significant genes in these tissues and/or other tissues,

XX and also permits the acquisition of molecular markers useful in breeding

XX schemes, genetic and molecular mapping, and in cloning of agronomically

XX significant genes. The nucleic acid molecules are further useful for

XX detecting the expression level or pattern of a protein or mRNA and for

XX detecting the presence or quantity of a protein by tissue printing. The

CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nucleon33B gynoecium tissue cDNA library (Lib3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ Sequence 534 BP; 181 A; 105 C; 115 G; 133 T; 0 U; 0 Other;

Query Match 15.0%; Score 217.8; DB 13; Length 534;
Best Local Similarity 65.2%; Pred. No. 6.8e-38;
Matches 337; Conservative 0; Mismatches 177; Indels 3; Gaps 1;
QY 542 TAGAACCCCAAGCAAAATACAGCATGACTACTACACAGTGCACAGAGTGG 601
Db 17 TTGACGACCCCAATCACTTACAGCATGATTTTATCAGAAACAGAGAGTGG 76
QY 602 TACTGACAAATTTGCTAAGGTGTTCTGCTGATAGTGTATGATTTTGTGTAAC 661
Db 77 TTGTCACAAATTTGCCAAGGAATACCAACGCGAGTGTAAAGTTAATTATGTTGAAC 136
QY 662 AGATGTTGAGTGTATCCATTTGAAGTTCTGTTGTAAGAACCATACCATTTTCAGCCCGTC 721
Db 137 AAATGCTAAGTGTCTGCTGATGCAATGACCGCAAGATGCTTATCATCTCCAACTCGCT 196
QY 722 TGTCTTAAAGATTATCCCTGAGAAATCAATATCAAGTCTTATCCACCAAGTTGAAA 781
Db 197 TATTTGGAAGATAATACCTGACAGTGCAGATATGATGTTTGTCAACCAAAATTGAAA 256
QY 782 TAGCCTTGCACAACTGAGCAGGTGACATGACAAACCTGATTTATAGTGAAGACCA 841
Db 257 TTAGTTAGCAAAAGCTGAACCAATTCATGAGCATCCCTCGAAATTTAGCATGGAAGTTG 316
QY 842 AGACTGTTCCCAAGAGTAAGCAGCGCAGCTGAAACAGCCCAAGACTTCATACCCAT 901
Db 317 CTGTGCCCAAGAGTAAATGATGTTCTGTTGCTGCAATCAAGACATGTTACCCAT 376
QY 902 CTTCAAAGCGGAAAA---AAGACTGGGATAAATCGGAAGCTGAAGTCAAAAGGAGGAGA 958
Db 377 CCTCAAACCAAAAGGGTTGATTTGGGATAAATCGAAGCTCAAGTGAAGAGGAGAGA 436
QY 959 AGGAGAAAATCTGATGTTGATGCTGATGCAATGAACAAATTTCTCGTGACATCTCAAGG 1018
Db 437 AAGATGAAAGCTAGATGTTGATGAGCTTTGAATAAATTTTTCGCGACATTTATCAGG 496
QY 1019 ATGCTGATGAAGATGCGGAGGCGCATGATGAAGTC 1055
Db 497 ATGCTGACGAGGATCCAGAGAGGCGCATGCAGAAATC 533

RESULT 14
ABL75330
ID ABL75330 standard; cDNA; 272 BP.

XX
XX
AC ABL75330;
XX

DT 14-MAY-2002 (first entry)

DE Corn tassal-derived polynucleotide (cdps) SEQ ID NO:4704.

XX Corn; corn tassal-derived polynucleotide; cdps; hybrid breeding; CDPS;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassal; gene; ss.

XX Zea mays.

XX US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-00294093.

XX 21-APR-1998; 98US-0082567P.

XX (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
XX WPI; 2002-163647/21.

XX Novel purified corn tassal-derived polynucleotide useful for determining
PT altered gene expression, to recover regulatory elements and to follow
PT inheritance of desirable characteristics through hybrid breeding
PT programs.

PS Claim 1; SEQ ID NO 4704; 201pp; English.

CC The present sequence describes a purified corn tassal-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassal-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful in
CC the evaluation, and alteration of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC multigene traits in a plant breeding program. (I) can be used to produce
CC a tassal-specific profile of gene transcription, a transcript image, to
CC clone regulatory elements for use in transformation vectors, to express a
CC polypeptide, to identify, isolate or extend identical or related corn
CC tassal nucleic acid sequences from DNA libraries, in nucleic acid
CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences

SQ Sequence 272 BP; 95 A; 54 C; 70 G; 49 T; 0 U; 4 Other;

Query Match 14.9%; Score 216.4; DB 6; Length 272;
Best Local Similarity 90.9%; Pred. No. 1.1e-37;

Matches 241; Conservative 0; Mismatches 19; Indels 5; Gaps 1;

QY 784 CGCTTGCACAAAGCTGAGCAGTGCATGACATGGACACACCTGGATTATAGTGAAGACCAAG 843
Db 1 CGCTTGCACAAAGCTGAGCAGTGCATGACATGGACACACCTGGATTATAGTGAAGACCAAG 60

QY 844 ACTGTTCCCAAGAGATAAGCAACCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCT 903
Db 61 GCTATTCCCAAGAGATAAGCAACCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCT 115

QY 904 TCAAGGCGAAAAAGACTGGGATAAAGTGAAGTCAAGTCAAAAGAGGAGGAGGAGAA 963
Db 116 TCAAGTCAAAAGAGGACTGGGATAAAGTGAAGTCAAGTCAAAAGAGGAGGAGGAGAA 175

QY 964 GAAAAAATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1023
Db 176 GAAAAAATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 235

QY 1024 GATGAAGATATGCGGAGGCGCATCA 1048

Db 236 GATGAAGATATGCGGAGGCGCATCA 260

RESULT 15

ACL25189/c

ID ACL25189 standard; DNA; 541 BP.

XX ACL25189;

XX AC

DT 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

This Page Blank (uspto)

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	164.4	11.3	272	4	US-09-313-294A-4573	Sequence 4573, Appl	
2	156.4	10.7	2053	3	US-09-227-357-45	Sequence 45, Appl	
3	76.2	5.2	198	4	US-09-248-796A-2330	Sequence 2330, Appl	
4	74.2	5.1	1735	4	US-09-949-016-1499	Sequence 1499, Appl	
5	74.2	5.1	1982	4	US-09-016-434-1067	Sequence 1067, Appl	
6	74.2	5.1	1982	4	US-09-825-497A-40	Sequence 40, Appl	
7	74.2	5.1	1987	4	US-09-825-497A-39	Sequence 39, Appl	
8	74.2	5.1	1987	4	US-09-517-779-1	Sequence 1, Appl	
9	73.2	5.0	425	4	US-09-513-999C-4027	Sequence 4027, Appl	
10	63.4	4.4	47493	4	US-09-949-016-13241	Sequence 13241, A	
11	62.4	4.3	2735	4	US-09-551-974A-101	Sequence 101, Appl	
12	62.4	4.3	2735	4	US-09-565-501A-101	Sequence 101, Appl	
13	62.4	4.3	2735	4	US-09-639-206A-101	Sequence 101, Appl	
14	62.4	4.3	2735	4	US-09-874-923-101	Sequence 101, Appl	
15	62.4	4.3	3012	4	US-09-551-974A-94	Sequence 94, Appl	
16	62.4	4.3	3012	4	US-09-565-501A-94	Sequence 94, Appl	
17	62.4	4.3	3012	4	US-09-639-206A-94	Sequence 94, Appl	
18	62.4	4.3	3012	4	US-09-874-923-94	Sequence 94, Appl	
19	62.4	4.3	3134	2	US-08-533-669A-1	Sequence 1, Appl	
20	62.4	4.3	3134	3	US-09-483-861-1	Sequence 1, Appl	
21	62.4	4.3	3134	3	US-09-022-765-1	Sequence 1, Appl	
22	62.4	4.3	3134	4	US-09-551-974A-1	Sequence 1, Appl	
23	62.4	4.3	3134	4	US-09-565-501A-1	Sequence 1, Appl	
24	62.4	4.3	3134	4	US-09-639-206A-1	Sequence 1, Appl	
25	62.4	4.3	3134	4	US-09-874-923-1	Sequence 1, Appl	
26	62.4	4.3	3134	4	US-08-798-841-1	Sequence 1, Appl	
27	62.4	4.3	4233	4	US-09-551-974A-99	Sequence 99, Appl	

RESULT 2
US-09-227-357-45
; Sequence 45, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010PI
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12

; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 2053
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-227-357-45

Query Match 10.7%; Score 156.4; DB 3; Length 2053;
Best Local Similarity 56.3%; Pred. No. 5.7e-29;
Matches 339; Conservative 0; Mismatches 251; Indels 12; Gaps 2;

Qy 560 CAAAATACAGCGATGACTACTCAACAGTGCACAGAGTGGTACTGACATAATTTGCTA 619
Db 1416 CAAAATCAAGTATGACTGGTATCAAAACAGANTCTCAAGTAGTCATACATTATGATCA 1475
Qy 620 AGGTGTTCTCTGCTGATAGTGTAGTCATTTGTTGGTGAACAGATGTTGAGTGATCCA 679
Db 1476 AGATGTTTCAAGAGATGATGTAATGTGGAATTTTCAGAAAAAGAGTCTCTGCTTTGG 1535
Qy 680 TTGAAGTTCTCTGGTGAAGAACCATACCATTTTCAGCCCCCTCTGTTTCTTAAGATTATCC 739
Db 1536 TTAACACTTCTCTGGAGAGGATTTACAAATTTGAAACTGGAACTTCITTCATCCTATAATAC 1595
Qy 740 CTGAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCTTGCAGAAAGCTG 799
Db 1596 CAGAACAGAGCAGCTTTAAAGTACTTTCAACAAGAGATTGAANTTAACCTGAAAAAGCCAG 1655
Qy 800 AGCAGGTGACATGACCAACCCCTGGATTTATAGTGAAGAACCAAGACTGTTCCCAAGAGA 859
Db 1656 AGGTGTTGAGATGGAAAAAGCTAGA-----GGGGCAAGGAGATGTGCTACGCCAAAAC 1709
Qy 860 TAACACGCCAGCTGAACAGCCCCCAAGAGCTTTCATACCCATCTTCAAGCGGAAAAAAG 919
Db 1710 AATTCGTAGCAGATGTAAAGAACCTATATCATCATCTCTCTT-----ATACAGAA 1763
Qy 920 ACTGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAGGAGAAAGAAAATCTGATGGTG 979
Db 1764 ATTGGATAAATTTGTTGGTGAGATCAAGAGAAAGAAAGAAATGAAAGTTGGAGGGAG 1823
Qy 980 ATGCTGCATTGAAACAAATTTCTTCGTGACATCTACAGAGATGCTGATGACAGATATCGGA 1039
Db 1824 ATGCAGCTTTAAACAGATTTATTTTCAGCAGATCTATTTCAGATGGTTCTGATGAAGTGAAC 1883
Qy 1040 GGGCCATGATGAAGTCTATCTGTTGGAATCAAAATGGCACTGTTCTCTCAACCAATTTGGAAG 1099
Db 1884 GTGCCATGAACAAATCTTTTATGAGTCCGGTGGTACAGTTTTCAGTACCACTGTTCTG 1943
Qy 1100 ATGTTGGAGCAAGAGAGGTAGAGGGAGCCCTGATGTTGTTATGAGTCAAGAGTGGG 1159
Db 1944 ATGATAGTAAAGGAAAGTTGAAATCAATCTCTCTGATGATATGGAATGGAAGAAAGTACT 2003
Qy 1160 AA 1161
Db 2004 AA 2005

RESULT 3
US-09-248-796A-2330
; Sequence 2330, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725

;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 2330
;; LENGTH: 198
;; TYPE: DNA
;; ORGANISM: Candida albicans
US-09-248-796A-2330

Query Match 5.2%; Score 76.2; DB 4; Length 198;
Best Local Similarity 64.4%; Pred. No. 2.2e-09;
Matches 114; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 991 AACAAATTTCTCGTGACATCTCAAGGATGCTGATGAAGATATCGGAGGCGCATGATG 1050
DB 22 AATGACTTTTCCGGAATATTTCAAGATGTTGACGAAGATTTAGAAAGCGATGATG 81
QY 1051 AAGTCATTCGTGGAATCAATGCGACTGTTCTCAACCAATTTGGAAGATGTTGGAGCA 1110
DB 82 AAGAGTTATGTTCAATCGAATGGAATGTTTGTACTACAAATTTGGATGAAGCTAAAGAT 141
QY 1111 AAGAAGGTAGAAGGAGCGCCCTCATGATGATGAGCTCAAGAAGTGGCAATACTAA 1167
DB 142 AAGAATTGAGGTTTACCACCAAGATGGATGGAAGTCAAGAAATGGGATACGTAA 198

RESULT 4
US-09-949-016-1499
;; Sequence 1499, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1499
;; LENGTH: 1735
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-1499

Query Match 5.1%; Score 74.2; DB 4; Length 1735;
Best Local Similarity 55.1%; Pred. No. 1.9e-08;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCTTTCGTGCGACGAGACTTCGAGCTGGCC 147
DB 82 CGAGAGGAGCTCAAGACTCAGGCCAATGACTTCTCAAAGCCAAAGGACTACGAGAACGCC 141
QY 148 ACCGAGCTCTACAGCCAGGCCATCGAGCGCGCGCCGACCGCCGACCTCTATGCGGAC 207
DB 142 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCGAGCAATGCCATCTACTATGGCAAC 201
QY 208 CGCGCCCGAGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
DB 202 CGAGGCTGGCTTACCTGCGCACTAGTGTATGGCTACGCTGGAGAGCGCCAGCGG 261
QY 268 GCAATTGAGCTTGATCTTATGATGATAAGCTTACTACCGGAAAGGTCTGTCATGCAATT 327
DB 262 GCCATTGAGCTGACAAAGAAGTACATCAAGGGTTATTACCGCGCGGCTGCCAGCAACATG 321
QY 328 AAGCTTGAAGATACCAAACTGC 350

DB 322 GCACTGGGCAAGTTCCGGGCGGC 344
RESULT 5
US-09-016-434-1067
;; Sequence 1067, Application US/09016434
;; Patent No. 6500938
;; GENERAL INFORMATION:
;; APPLICANT: Janice Au-Young
;; APPLICANT: Jeffrey J. Seilhamer
;; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;; NUMBER OF SEQUENCES: 1490
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;; STREET: 3174 PORTER DRIVE
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;; CURRENT APPLICATION DATA:
;; FILING DATE: HEREWITH
;; APPLICATION NUMBER: US/09/016,434
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Zeller, Karen J.
;; REGISTRATION NUMBER: 37,071
;; REFERENCE/DOCKET NUMBER: PA-0002 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650) 855-0555
;; TELEFAX: (650) 845-4166
;; INFORMATION FOR SEQ ID NO: 1067:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1982 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GENBANK
;; CLONE: g1122930
US-09-016-434-1067

Query Match 5.1%; Score 74.2; DB 4; Length 1982;
Best Local Similarity 55.1%; Pred. No. 2e-08;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCTTTCGTGCGACGAGACTTCGAGCTGGCC 147
DB 60 CGAGAGGAGCTCAAGACTCAGGCCAATGACTTCTCAAAGCCAAAGGACTACGAGAACGCC 119
QY 148 ACCGAGCTCTACAGCCAGGCCATCGAGCGCGCGCCGACCGCCGACCTCTATGCGGAC 207
DB 120 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCGAGCAATGCCATCTACTATGGCAAC 179
QY 208 CGCGCCCGAGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
DB 180 CGAGGCTGGCTTACCTGCGCACTAGTGTATGGCTACGCTGGAGAGCGCCAGCGG 239
QY 268 GCAATTGAGCTTGATCTTATGATGATAAGCTTACTACCGGAAAGGTCTGTCATGCAATT 327
DB 240 GCCATTGAGCTGACAAAGAAGTACATCAAGGGTTATTACCGCGCGGCTGCCAGCAACATG 299
QY 328 AAGCTTGAAGATACCAAACTGC 350


```
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4027
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 46..417
US-09-513-999C-4027

Query Match      5.0%; Score 73.2; DB 4; Length 425;
Best Local Similarity 57.4%; Pred. No. 1.8e-08;
Matches 132; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY      88 GCGTCGATCTGAGAGCAAGCCAGGAGGCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db      127 GAGAGGAGCTCAAGACTCAGGCCATGACTACTTCAAGCCAGGACTACGAGACGCC 186

QY      148 ACCGAGCTCTACAGCCAGCCATCGACGCGCGCCGCCACCCGCGACCTCTATGCGGAC 207
Db      187 ATCAAGTTCTACAGCCAGCCATCGAGCTGAACCCAGCAATGCTACTTACTATGCGAAC 246

QY      208 CCGGCCAGCGGCACATCAAGCTCGGCAACTACATCAAGCTGTGGCGGATGTAACAAA 267
Db      247 CGAGCGCTGGCTTACCTGCGCACTGAGTGTATGCTACGCGTGGGAGACGCCAGCGG 306

QY      268 GCAATTGAGCTTGATCTCTATGATGATCAAAAGCTTACTACGGAAAGGTGC 317
Db      307 GCAATTGAGCTGCAAGAAGTACATCAAGGGTTATTACCGGGGCTGC 356

RESULT 10
US-09-949-016-13241
; Sequence 13241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13241
; LENGTH: 4793
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13241

Query Match      4.4%; Score 63.4; DB 4; Length 4793;
Best Local Similarity 54.5%; Pred. No. 4.6e-05;
Matches 127; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY      118 GCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCGCATCGACGCC 177
Db      8641 GTCTCGGAGCCAGGACTACGAGACGCGTCATCAAGTTCTACAGCCAGGCGCATCGAGCTG 8700
```

```
QY      178 GGGCCGCGCACCGCGACCTCTATCGACCGCGCCAGCGCACATCAAGCTCGGCAAC 237
Db      8701 AACCCAGCATGCCATCTACTATGGCAACCGAGCTGGCTACTCTGGCAGTGTGTCG 8760

QY      238 TACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGAATCTATGATGCAATAA 297
Db      8761 TATGGCTAGCGCTGGGAGACGCCAGCGGGCCATTGAGCTGGACAAGAGTACATCAAG 8820

QY      298 GCTTACTACCGAAAGTGCTGCTATGCTTAAGCTTGAAGATATACCAAACTGC 350
Db      8821 GGTATTATACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8873

RESULT 11
US-09-551-974A-101
; Sequence 101, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551,974A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; comprising multiple Leishmania antigens
US-09-551-974A-101

Query Match      4.3%; Score 62.4; DB 4; Length 2735;
Best Local Similarity 51.4%; Pred. No. 2.1e-05;
Matches 144; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY      65 AAGCCGGAAGGGGCGCCATGCGCGCTGCGATCTGAGAGCAAGCGCAAGGAGGCTTCG 124
Db      1676 AAGCCCTACATCGATCTGAGATCGCAAGCAAGAAAGCAAGGTAACCAAGTACTTCA 1735

QY      125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCGCATCGCGGCGCCG 184
Db      1736 AGGAGATAGTTTCCCGAGGCGCTGCGAGCGTACAGGAGGCCATCAAGCGCAACCTG 1795

QY      185 CCACCGCGGACCTCTATGCGGACCGCGCCAGCGCACATCAAGCTCGGCAACTACACTG 244
Db      1796 CCGAGCACACCTCTACAGCAATCGCGCGCGGTACATCAAGCTTGGAGCTTCAACG 1855

QY      245 AGCTGTGGCGGATGCTAACAAGCAATTGAGCTTGAATCTATGATGCAATAAGCTTACT 304
Db      1856 ACGCCCTCAAGGACGCGGAGAGTGCATTGAGCTGAAGCCCGCATTTGTTAAGGGTACG 1915

QY      305 ACCGGAAAGGTGCTGCATGCTTAAAGCTTGAAGATATACCA 344
Db      1916 CGCGCAAGGGTCTGCTTACTTTTGGACCAAGCAGTACAA 1955

RESULT 12
US-09-565-501A-101
; Sequence 101, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
```

```

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565,501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-565-501A-101

Query Match      4.3%; Score 62.4; DB 4; Length 2735;
Best Local Similarity 51.4%; Pred. No. 2.1e-05;
Matches 144; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 65 AAGCCGGAAGGGCGCCATGCGCGTCTGAGATCTGAGAGCAAGGCCAAGAGGCGCTTCG 124
Db      |||||
QY 125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCAGGCCATCGACCGCGGCCG 184
Db      |||||
QY 1736 AGGAGGATAAGTTCCCGAGGCGCTGCGAGCGGATACCGAGGCCATCAAGCGCAACCTG 1795
Db      |||||
QY 185 CCACCGCCGACCTCTATGCGACCGCGCCCGCCAGCGGCACATCAAGCTCGGCAACTACTG 244
Db      |||||
QY 1796 CCGAGCACACCTCTACAGCAATCGCGCGCGGTACATCAAGCTTGAGGCGCTTCAAG 1855
Db      |||||
QY 245 AGCGTGTGGCGGATGTAACAAGCAATTGAGCTTGATCTTATGATGATGATGATGATGAT 304
Db      |||||
QY 1856 ACGCCCTCAAGGACGCGGAGAGTGCATTGAGCTGAAGCGCGACTTTGTTAAGGGCTAG 1915
Db      |||||

RESULT 13
US-09-639-206A-101
; Sequence 101, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-639-206A-101

Query Match      4.3%; Score 62.4; DB 4; Length 2735;
Best Local Similarity 51.4%; Pred. No. 2.1e-05;
Matches 144; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 65 AAGCCGGAAGGGCGCCATGCGCGTCTGAGATCTGAGAGCAAGGCCAAGAGGCGCTTCG 124
Db      |||||
QY 125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCAGGCCATCGACCGCGGCCG 184
Db      |||||
QY 1736 AGGAGGATAAGTTCCCGAGGCGCTGCGAGCGGATACCGAGGCCATCAAGCGCAACCTG 1795
Db      |||||
QY 185 CCACCGCCGACCTCTATGCGACCGCGCCCGCCAGCGGCACATCAAGCTCGGCAACTACTG 244
Db      |||||
QY 1796 CCGAGCACACCTCTACAGCAATCGCGCGCGGTACATCAAGCTTGAGGCGCTTCAAG 1855
Db      |||||
QY 245 AGCGTGTGGCGGATGTAACAAGCAATTGAGCTTGATCTTATGATGATGATGATGATGAT 304
Db      |||||
QY 1856 ACGCCCTCAAGGACGCGGAGAGTGCATTGAGCTGAAGCGCGACTTTGTTAAGGGCTAG 1915
Db      |||||

RESULT 14
US-09-874-923-101
; Sequence 101, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-874-923-101

Query Match      4.3%; Score 62.4; DB 4; Length 2735;
Best Local Similarity 51.4%; Pred. No. 2.1e-05;
Matches 144; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 65 AAGCCGGAAGGGCGCCATGCGCGTCTGAGATCTGAGAGCAAGGCCAAGAGGCGCTTCG 124
Db      |||||
QY 1676 AAGCTTACATCGATCTCTGAGATCGGAGAGCAAGAGCAAGAGTAAACAGTACTTCA 1735
Db      |||||
QY 125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCAGGCCATCGACCGCGGCCG 184
Db      |||||
QY 1736 AGGAGGATAAGTTCCCGAGGCGCTGCGAGCGGATACCGAGGCCATCAAGCGCAACCTG 1795
Db      |||||
QY 185 CCACCGCCGACCTCTATGCGACCGCGCCCGCCAGCGGCACATCAAGCTCGGCAACTACTG 244
Db      |||||
QY 1796 CCGAGCACACCTCTACAGCAATCGCGCGCGGTACATCAAGCTTGAGGCGCTTCAAG 1855
Db      |||||
QY 245 AGCGTGTGGCGGATGTAACAAGCAATTGAGCTTGATCTTATGATGATGATGATGATGAT 304
Db      |||||
```

Db 1856 ACCCCCTCAAGGACGGGAGAGTGCATTGAGCTGAAGCCGACCTTGTTAAGGGCTACG 1915
QY 305 ACCGGAAGGTGCTGCATGCAATTAAGCTTGAAGAAATACCA 344
Db 1916 CGGCAAGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955

RESULT 15

US-09-551-974A-94
; Sequence 94, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yaghir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551,974A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 3012
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-551-974A-94

Query Match 4.3%; Score 62.4; DB 4; Length 3012;
Best Local Similarity 51.4%; Pred. No. 2.2e-05;
Matches 144; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
QY 65 AAGCGGAAGGGCGCCATGGCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCGCTTCG 124
Db 1676 AAGCCTACATCGATCTGAGATCGCGAAGCAGAAGAAAGCAAGAGTAAACCACTTCA 1735
QY 125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGAGCGGCGCG 184
Db 1736 AGGAGATTAAGTTCCCGAGGCGGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCTG 1795
QY 185 CCACCGCCGACCTCTATGCGCGCCAGCGCCACATCAAGCTCGGCAACTACACTG 244
Db 1796 CCGAGCACACCTCCTACAGCANTCGCGCGCGGTACATCAAGCTTGGAGCCTTCAAG 1855
QY 245 AGGCTGTGGGATGTACAAGCAATTGAGCTTGCATCTCTATGATGCATAAAGCTTACT 304
Db 1856 ACCCCCTCAAGGACGGGAGAGTGCATTGAGCTGAAGCCGACCTTGTTAAGGGCTACG 1915
QY 305 ACCGGAAGGTGCTGCATGCAATTAAGCTTGAAGAAATACCA 344
Db 1916 CGGCAAGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955

Search completed: April 23, 2005, 08:59:03
Job time : 261.537 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:35:27 ; Search time 6457.31 Seconds
(without alignments)
10918.214 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcgacgagcgccac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	750.4	51.6	1409	8 AF439974	AF439974 Hordeum v
2	743.6	51.1	1222	8 AF192467	AF192467 Oryza sat
3	456.2	31.4	984	8 RID251317	AJ251317 Rubus ida
4	454.8	31.3	1113	8 AF494083	AF494083 Nicotiana
5	420	28.9	1366	8 AY085464	AY085464 Arabidops
6	419	28.8	1354	8 ATH118019	AJ118019 Arabidops
7	418.4	28.8	1108	8 AY150487	AY150487 Arabidops
8	418.4	28.8	1290	8 AF439976	AF439976 Arabidops
9	418.4	28.8	1400	8 AF428340	AF428340 Arabidops
10	418.4	28.8	1420	8 AF370229	AF370229 Arabidops
11	417.6	28.7	1077	6 AX412470	AX412470 Sequence
12	413.8	28.4	1197	8 AJ620883	AJ620883 Brassica
13	407.2	28.0	1068	8 AJ620882	AJ620882 Brassica
14	394.8	27.1	1291	8 AF439975	AF439975 Arabidops
15	393.2	27.0	1084	8 AY096395	AY096395 Arabidops
16	393.2	27.0	1348	8 AY063978	AY063978 Arabidops
17	295.8	19.6	968	8 ATH118020	AJ118020 Arabidops
18	261	17.9	582	8 AK059268	AK059268 Oryza sat
19	207.8	14.3	634	8 AF516180	AF516180 Nicotiana

20	207.8	14.3	634	8	AF516181	Nicotiana
21	164.4	11.3	272	6	AR249214	Sequence
22	163	11.2	1535	9	BC000911	Homo sapi
23	163	11.2	1555	9	HS344097	Homo sapi
24	163	11.2	1776	6	CQ490463	Sequence
25	163	11.2	1776	6	CQ496307	Sequence
26	160	11.0	1606	5	AJ720706	Gallus ga
27	159.6	11.0	1518	5	BC072118	Xenopus l
28	159	10.9	1235	9	AY321358	Homo sapi
29	156.4	10.7	1002	6	CQ724202	Sequence
30	156.4	10.7	1002	9	AF132856	Homo sapi
31	156.4	10.7	1025	6	BD209970	Human tra
32	156.4	10.7	2053	6	BD192630	123 human
33	156.4	10.7	2053	6	AR184101	Sequence
34	156.2	10.7	1002	9	BT009798	Homo sapi
35	146.6	10.1	139201	8	AP002970	Oryza sat
36	146.6	10.1	150022	8	AP003213	Oryza sat
37	145.2	10.0	1314	5	CR761938	Xenopus t
38	143.6	9.9	153865	8	AP002744	Oryza sat
39	139.4	9.6	1104	10	BC009167	Mus muscu
40	119	8.2	1089	9	AF068289	Homo sapi
41	114.6	7.9	154558	8	AC107206	Oryza sat
42	111	7.6	137843	9	AC138972	Homo sapi
43	107.8	7.4	304	4	BT344098	Bos tauru
44	92.2	6.3	1293	8	AK073324	Oryza sat
45	90.8	6.2	1242	6	AX488766	Sequence

ALIGNMENTS

RESULT 1
AF439974
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

AF439974 1409 bp mRNA linear PLN 06-NOV-2003
Hordeum vulgare SGT1 (SGT1) mRNA, complete cds.
AF439974.1 GI:17017305
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 1409)
Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A.,
Shirasu, K. and Schulze-Lefert, P.
The RAR1 interactor SGT1, an essential component of R
gene-triggered disease resistance
Science 295 (5562), 2073-2076 (2002)
21893744
11847307
2 (bases 1 to 1409)
Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A.,
Shirasu, K. and Schulze-Lefert, P.
Direct Submission
Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre,
Colney Lane, Norwich NR47UH, UK
Location/Qualifiers
1. .1409
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
1. .1409
/gene="SGT1"
9. .1130
/gene="SGT1"
/codon_start=1
/product="SGT1"
/protein_id="AAL33610.1"
/db_xref="GI:17017306"
translation="MAAANASDLESKAEAFVDDDFELAELVTOAEAGPATAELVA
DRAQHIKLGSTVEAADANKAIEIDPSMHKAYLRKGSACIKLEEYKAAALVSGSS
YASGDSRFRILMECCDDRIAEASQAPVRNAAANAVPATSSGATTVVTEADQDGENM

ENAOPTVEVPSKPKYRHDYNTPTPEVLITFAKGVPADSVVVDFGQMLSVSIELPGE
EYPHFQPRLESKIVDPDKCYTVLSTKVEIRLKAEPYTWTSJDYTGKPAQKINVPA
ESASRPSYPSKSKDWDKLEAEVYQSKDEKLJDAAALNFFREIYSDADEDMRAM
MKSPFVEJNPTLSTNWKQCTKVEGSPDPGSMELKKWYE"

ORIGIN	Query Match	51.6%	Score 750.4	DB 8	Length 1409
	Best Local Similarity	79.5%	Pred. No. 1.7e-120		
	Matches 927	Conservative	0	Mismatches 206	Indels 33
				Gaps	
QY	69	CGGAAGGGCGCCATGGCCGCGTCGGATCTCGAGAGCAAGGCCAAGAGAGCCCTTCGTCGA	128		
DB	5	CAGATGGCGCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGAGAGCCCTTCGTCGA	64		
QY	129	CGAGCACTTCGAGTGGCCACCGAGCTCTACAGCCAGGCCATCGACCCGGCGCCGCCAC	188		
DB	65	CGAGCACTTCGAGTGGCCACCGAGCTCTACACCAGGCCATCGAGCTGGCCAGCCAC	124		
QY	189	CGCCGACCTCTATGCCGACCGCGCCAGGGGCAATCAAGCTCGGCCAACTACACTGAGGC	248		
DB	125	CGCCGAGCTCTACGCCGACCGGCCAGGCTCACATCAAGCTGGCGCAGTTACACTGAGGC	184		
QY	249	TGTGGCGGATGCTAAACAAAGCAATGAGCTTGATCTCTATGATCATAAAGCTTACTACCG	308		
DB	185	TGTAGCTGATGCCAACCAAGCAATGGAATCTTGATCTCTCGATGCAAAAGCGTACTTCG	244		
QY	309	GAAAGGTGCTCATGCTAATGAAGCTTGGAAGATACAAACTGCAAGGCTGCTCTTGAGTT	368		
DB	245	GAAGGTTCTGCTTGATCAAGCTCGAGGATACCAACTGCAAGGCTGCTCTTGAGTT	304		
QY	369	GGGTCTCTTATGATCAGCGGATCAAGGTTTGCTCGTCTATTTGAAGGAATGTGATGA	428		
DB	305	TGGTTCTCTTATGATCAGCTGAGTCAAGGTTTACTCGTCTTATGAAGGAGTGTGATGA	364		
QY	429	GCGCATCGCTGAGGAATCTAGCCAGGCCACCAAGTAAAGGAATGTGAGGCTACTGTGGCT	486		
DB	365	TGTTATTCGCTGAGGAGGCTAGCCAGCGCCAGTAAAGAAATGCGCTGCGGCTGTTCCTCC	424		
QY	487	-----GCTACTATTGAGNCAAGGAGGATTTCCACAAA	518		
DB	425	TGCTACATCTTCGCGGGCTACAACTGTGGTTTACTGAAGCTGAGGACCAAGATGGTGA	484		
QY	519	TATGGAGAAATACACCAAGTATAGAACCCCAAGCAAAACAAATACAGCATGACTA	578		
DB	485	TATGGAGAAATACACAGCCAAACGTTAGAGTGCCAAAGCCCAAAATACAGCATGACTA	544		
QY	579	CTAACACAGTGCCACAGAAGTGTTACTGACATATTTGCTAAGGTTGTTCTTCCTGATG	638		
DB	545	CTACAATCTCTCAGAAAGTGTTACTGACTATATTTGCTAAGGTTGTTCTTCAGCTGAC	604		
QY	639	TGTAGTCAATTGTTGGTGAACAGATGTTGAGTGTATCCATGAAAGTTCTTGGTGAAGA	698		
DB	605	CGTGGTTGTTGACTTTGGTGAACAGATGCTGAGTGTCTCAAATGAACTTCTTGGTGAGA	664		
QY	699	ACATACCAATTTTCAGCCCGCTGTTTCTAAGATATCCCTGAGAAATGCAATATCA	758		
DB	665	ACCATACCAATTTTCAGCCCGCTGTTTCTAAGATATCCCTGAGAAATGCAATATCA	724		
QY	759	AGTCTTTATCCACCAAGGTTGAAATACGCTTTCGAAAGGCTGAGCAGGTGACATGGACA	818		
DB	725	TGTTATGTTCTACAAAGGTCGAAATACGCTTTCGAAAGGCTGAGCAGGTGACATGGACA	784		
QY	819	CTTGGATATATAGTGGAAAGCAAGACTGTTTCCCGAAGATGAAGCAGCCAGCTCAAC	878		
DB	785	ACTGGATATATAGTGGTAAACCAAG---GCTCTCTCAGAGATGAATGTTACCACTGATC	841		
QY	879	AGCCCCAAGACCTTCATACCCATCTTCAAAGGGCGAAAAAGACTGGGATAAACTGGAAG	938		
DB	842	AGCCCCAAGACCTTCATCTTCAATCAAAATCCAAAAAGCACTGGGATTAAGCTTGAGGC	901		
QY	939	TGAAGTCAAAAGGAGGAGGAAGAAACCTTGATGGTGTGCTGCTCATTTGAACAAATT	998		
DB	902	TGAAGTGAATAAAGCAGGAGAGGATGAGAACTTGAACGTTGCTGCTGCTGATGAAACAGTT	961		

```

103 1118 CCACCTGATGGATCGAGCTTAAGAAATGGAGTACTAAGATCTCGTCTCGTCCATGC 1177
1188 TGTAATTC 1196
1178 TGTAATTC 1186

RESULT 3
RID251317
LOCUS
DEFINITION
Rubus idaeus partial mRNA for putative protein phosphatase.
ACCESSION
AJ251317.1 GI:6468694
VERSION
protein phosphatase.
KEYWORDS
Rubus idaeus
ORGANISM
Rubus idaeus
Vikaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rubus.

REFERENCE
1 Schroeder, G., Lurz, G. and Schroeder, J.
cDNA for protein phosphatase like protein from raspberry (Rubus
idaeus)
Unpublished
2 (bases 1 to 984)
Schroeder, G.
Direct Submission
JOURNAL
TITLE
Submitted (06-NOV-1999) Schroeder G., Institut fuer Biologie II,
Universitaet Freiburg, Schaenzlestr. 1, Freiburg D-79104, Germany
Location/Qualifiers
FEATURES
source
1..984
/organism="Rubus idaeus"
/mol_type="mRNA"
/db_xref="taxon:32247"
<1..984
/codon_start=1
/product="putative protein phosphatase"
/db_xref="GI:6468694"
/db_xref="UniProt/TrEMBL:Q9SM53"
/translations="PKNSLYSDRAQNMKSNLLEAVADANKAIELDPSLFKAYLRK
GIACMLKEVYRTAKALELGLASLPGESRPFAKMIKECDLIAEENGPPKQPMETTTT
EIVAEDEVDPVPSNEVTVAPVKYHPEVQKAEVWVTFIAKGI PAENVAVDFGQ
ILSVSDVGEVHIHFQPLRFGKLIPEKCRFVLSTKVEIRLAKAENHHTSLFSKD
NPVLKNAPVIGAKPSYSPSKRPDMDKLEAQVKBEKLDGDAALNKFQDI
YKDAEDTTRAMEKSFVESNGTVLSTNWKEVGNKVEGSA PDGMEMKMKWEI"

Query Match 31.4%; Score 456.2; DB 8; Length 984;
Best Local Similarity 68.5%; Pred. No. 2.4e-69;
Matches 677; Conservative 0; Mismatches 303; Indels 9; Gaps 3;

QY 181 CCCGCCACCCGACCTCTATGCGACCGCGCCAGGCACATCAAGCTCGGCAACTAC 240
DB 1 CCCAAGACTCCGAGCTCTACTCGACCGGCTCAGCCCAACATGAATCCAAATCTC 60
QY 241 ACTGAGCTGTGGCGATGCTAACAAAGCAATTGAGCTTGATCTATGATGATAAGCT 300
DB 61 ACCGAGCTGTGGCGATGCGAACAGGCCATTGAGTTGGACCGCTCTTTGTTCAAAGCC 120
QY 301 TACTACGGAAAGGTGCTGCATGCATTAGCTTGAAGATACCAAACTGCAAGGCTGCT 360
DB 121 TACTTCGCAAGGAATTCGCTGATGAAGCTGAGGAATATAGAACTGCAAGGCGAGCC 180
QY 361 CTTGAGTTGGGTTCTTCTATGATCAGGGGATTCAGGTTTGCTCGCTATTCAAGGAA 420
DB 181 TTGGAATCTGGTCTCTTCTGGCACCGGGGAATCCAGATTCGTTAAATGATTAAGAG 240
QY 421 TGTGATGAGCGCATCGCTGAGGAATCTAGCAGGCAACCAAGTAAAGATGTTGAGGCTACT 480
DB 241 TGTGATGAGCTCATTCAGAGGAAATGGTGAGCCACC---GAAGCAACCAATGGAAACA 297
QY 481 GTGGCTGCTACTATTATGGGCAAGGAGGATTTCAAAATATGGGAATATACACCAGCTG 540
```

```

44 CCAAGCGCCATGCGACACCGCGCGCGTGGATCTGGAGAGCAAGCGCGGCGCTT 103
123 CGTCGACGACACTTCGAGCTGCGCCACCGAGCTCTACAGCAGGCCCATCGAOGCCGGGCC 182
104 CGTCGACGACGACTTCGAGCTCGCGCGAGCTCTACACGAGGCAATCGAGGCGAGGCC 163
183 CGCCACCGCGACCTCTATCGGACCGCGCGCGCGGCGCAGCATCAAGCTCGGCAACTACAC 242
164 CGCCACCGCGAGCTCTACGCGGACCGCGCGCGCGCCATATCAAGCTAGGCAACTACAC 223
243 TGAGGCTGTGGCGATGCTAACAAAGCAATAGCTTGAATCTATGATGATAAAGCTTA 302
224 TGAGGCTGTAGCTGATGCTAACAGGCGCATGAACTTGACCCCATGACACAGGCTTA 283
303 CTACCGGAAAGGTGCTGCATGCAATTAAGCTTGAAGATAACAACTGCAAGGCTGCTCT 362
284 TCTTCGTAAGAGGCGCTGCATGTATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCT 343
363 TGAGTTGGGTTCTTCTTATGCATCAGCGGATTCAGGTTTGGCTGCTATTTGAGGATG 422
344 TGAATGGGTTACTCGTTCGATCTGGTGAATCAAGGTTTACTCGCTTAATGAAGGAGTG 403
423 TGATGAGCGCATCGCTGAGGAATCTAGCAGGACCAAGTAAAGATGTTGAGG----- 475
404 TGATGAGCGCATGCTGAGGAGCTTACTGAATCCCTGTTAAGAGGCTGAGATGAGGC 463
476 -----CTACTGTGCTGCTACTATTTGAGGACCAAGGAGATTTCACAAATATGGAGAA 527
464 AGCTGCCCTCTGTTCTTCTTTGTTGAGGAAAGGATGATGCTGCAAAACATGGATAA 523
528 TACACCAAGGATGATAGACCCCAAGCAACCAAAATACAGGATGACTACTACAACAG 587
524 TACACCAACCAATGGTAGA-----AGTGAAGCCAAAATAACAGGCACGACTTCTACACAG 577
588 TGCCACAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 647
578 TGCTACAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 637
648 TGATTTGGTGAACAGATGTTGAGTGTATCATTTGAAGTTCTTGTTGAGGAGCAACATACCA 707
638 TGATTTGGTGAACAAATGTTAAGTGTGTCGATTTGAAGTCCCTTGAGAGGCGGTACCA 697
708 TTTTCAGCCCGCTGTTTCTTAAGATATCCCTGAGAAATGCAAAATATCAAGTCTTATC 767
698 TTTTCAGCCCTCGTCTGTTTCTTAAGATCATCCCTGAGAAAGCAGATACCAAGTGTATC 757
768 CACCAAGGTTGAATACGCTTGCACAAAGCTGAGCAGTGCATGAGCAACCCCTGGATTA 827
758 CACGAAGTTGAATAGAGCTGGCTAAGCTGAAACAGATTTACATGGACCTCATTGATTA 817
828 TAGTGAAGACCAAGACTGTTCCCAAGAGATAAGCAGCCAGCTGAAACAGCCCCAAG 887
818 TGATAAAAAACCAAGGCTGTTCCACAAAAGATAATCCCTCCAGCTGAATCGGCCAGAG 877
888 ACCTTCATACCATCTTCAAAGGGGAAAAAGACTGGGATAAATGGAAGCTGAAGTCAAA 947
878 GCCATCATATCTTCTTCAAATTCAAAGAAAGACTGGGATAAATGGAAGCTGAAGTTAA 937
948 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1007
938 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 997
1008 CATCTACAAGGATGCTGATGAAGATATCGGAGGCGCATGATGAAGTCAATTCGTTGAATC 1067
998 CATCTACAGTGTGATGATGAGACATCGCAGGAGCAATGATGAATCTTTTGTGATC 1057
1068 AATGGCACTGTTCTTCAAACCAATGGAAGATGTTGGAGCAAGAAAGGTAGAGGGAG 1127
1058 TAACGGTACTGTTCTGTCGACCAATTCGAAAGATGTTGGCTCGAAGAAAGGTAGAGGAG 1117
1128 CCCCCCTGATGGTATGAGCTCAAGAGTGGGAATACTAAAGTTTGAAGTCCCGCTCTT 1187
```



```

Db      298  ACAAATACAGAGATTGTCAGAGAGATTGTCAGCAGTGACCCGCTTCCATGAGGTG 357
Qy      541  ATAGAACCCCAAGCAACAAATACAGCAGTACATGCTACCAACAGTGCACAGAAAGTG 600
Db      358  ACAGTAGCAGCTGTCAAAACCAAAATACAGCAGTAACTTACCAAGAGCTGAAGAGTG 417
Qy      601  GTACTGCAATATTGCTAAGGTGTTCCGCTGATAGTGTAGTCAATGATTTGTTGAA 660
Db      418  GTTGTGACTATATTGCAAAAGGCAATACCTGCCGAAATGTTCTGTGATTTGTTCCA 477
Qy      661  CAGATGTTGAGTGATCCATTGAAGTTCCTGGTGAAGAACCAATACCATTTCAGCCCGT 720
Db      478  CAATATTAAGTGTCAGCATGATGTTCCGGTGAAGTACATATACATATTCAGCCTCG 537
Qy      721  CTGTTTCTAAGATTATCCCTGAGAAATGCAATATCAAGTCTTATCCCAAGGTTGAA 780
Db      538  CTATTTGAAAGATCAATCTCTGAAAGTGAGATTTGAGGTTTGTCTACCAAGTTGAA 597
Qy      781  ATAGCCCTTGCNAAAGCTGACAGGTGACATGGACCAACCTGGATTATAGTGAAGACCA 840
Db      598  ATTCCGCTTGCAAAAGCTGAACCTAATCACTGGACTTCTCTGAGTTCA---GCAAGGAT 654
Qy      841  AAGACTGTTCCCAAGAGATAAGCAGCAGCTGAAACAGCCCAAGACCTTCATACCCA 900
Db      655  AACCTGTCTCTAAGGTTAATGCCCCAGTTATTGGAGCCCAAGGCCATCTTACCCA 714
Qy      901  TCITCAAAGGGGAAAA---AAGACTGGGNTAACTGAAGCTGAAGTCAAAAAGGGAG 957
Db      715  TCCTCCAAACCAAAAGGGTAGATTGGGACAAGCTTGAAGCCCAAGTGAAGAGGAGAA 774
Qy      958  AAGCAAGAAACCTTGATGTTGATGCTGCAATGAACAAATTTCTCCCTGACATCTACAAG 1017
Db      775  AAGATGAGAAGCTTGATGATGATGAGCTTGAACAAATTTTCCAAGTATATACAAG 834
Qy      1018  GATGCTGATGAAGATATCCGAGGGCCATGATGAAGTCAATTCGTGGAATCAATGGCACT 1077
Db      835  GATGCTGATGAAGACACAAGAAAGGCCATGAGAAATCTTTTGTGAGTCAAAATGGGACG 894
Qy      1078  GTTCTCTCAACCAATTGAAGAGTTGGACCAAGAGGTAGAAGGGAGCCCTCTGAT 1137
Db      895  GTGCTTTCAAACAACTGGAAGAGTGGGCAAGAGGTGAGGAGGAGTGCCTCTGAT 954
Qy      1138  GGTATGGAGCTCAAGAAGTGGGAATACTA 1166
Db      955  GGTATGGAGATGAAGAAATGGAGATCTA 983

RESULT 4
LOCUS   AF494083
DEFINITION Nicotiana benthamiana 1113 bp mRNA linear PLN 30-OCT-2003
ACCESSION AF494083
VERSION   AF494083.1 GI:29468338
KEYWORDS
SOURCE   Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          asterids; lamiales; Solanales; Solanaceae; Nicotiana.
          1 (bases 1 to 1113)
REFERENCE Liu,Y., Schiff,M., Serino,G., Deng,X.W. and Dinesh-Kumar,S.P.
AUTHORS Role of SCF ubiquitin-ligase and the COP9 signalosome in the N
TITLE Plant-mediated resistance response to Tobacco mosaic virus
JOURNAL Plant Cell 14 (7), 1483-1496 (2002)
MEDLINE 22115173
PUBMED 12119369
REFERENCE 2 (bases 1 to 1113)
AUTHORS Liu,Y., Schiff,M., Serino,G., Dinesh-Kumar,S.P. and Deng,X.W.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2002) MCDB, Yale University, New Haven, CT 06520,
          USA
FEATURES
          Location/Qualifiers

```

```

source 1..1113
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
1..1113
/codon_start=1
/product="SGT1"
/protein_id="AA085509.1"
/db_xref="GI:29468338"
/translation="MASPLETRAKEAFIDDFELAVLDLYTOAIAMTPKNELFADRAQ
ANIKLNYTEAVDANKAELDPSMKAYLFGKGLACMKLEVOYOTAKALETGASLAPA
ESRFTLKICEDERIAEBEGLPNSVDKTSNVVAPPASESLONVAVAPKDAOPTVN
LSYOGSAAAPKPYRHEFYOKPEEVVVTIFAKGIPAKNVVDFGEQLSVISDVPDETY
SPQPLFKITPAKRYEVMKLEIRLAKAEPLHWTLSLEYTRASAVVQRNVSSDAP
RPSYPSKLRHVDMDKLEABVKBEKDEKLGDALINKFRDIYKDADEDETRAMWKS
FVESNGTVLSTNWEVGTKKVEGSPDGMELMKWEI"

ORIGIN
Query Match 31.3%; Score 454.8; DB 8; Length 1113;
Best Local Similarity 65.6%; Pred. No. 4.2e-69;
Matches 729; Conservative 0; Mismatches 347; Indels 36; Gaps 3;

Qy      88  GGTGCGATCTGGAGCAAGGCCCAAGGAGGCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db      4  GGTGCGATCTGGAGCAAGGCCCAAGGAGGCTTCATGTCAGCACACTTTGAGCTCGCC 63
Qy      148  ACCGAGCTCTACAGCCAGGCCCATCGACGCCGCCGCCGCCGACCTCTATGCCGAC 207
Db      64  GTTGACCTTTACATCAAGCAATTGCCATGACTCTCTAAGAACGCTGAGCTTTTCGCCGAC 123
Qy      208  CGGCCCGAGCGACATCAAGCTCGGCACTACATGAGGCTGTCGCGGATGCTTAACAA 267
Db      124  CGTGCTCAGGCCCAACATCAACTCACTACTCTGAAGCTGTTGTTGATCGGAACAAG 183
Qy      268  GCAATTGAGCTTGATCCTATGATGATAAAGCTTACTACCGGAAAGGTGCTGCATGCAT 327
Db      184  GCATTTGAGTTAGTCCITTCATGTCMAAAGCATATTTGCGTAAGGGTTGCGCTGTATG 243
Qy      328  AAGCTTGAAGAATACAAACCTGCAAGGCTGCTCTTGAAGTGGGTTCTTCTTATGCATCA 387
Db      244  AAGCTTGAAGAGTACCAAACTGCAAAAGCAGCTTTGGAAACTGGTGCTAGTTAGCACCG 303
Qy      388  GCGGATTCAGGTTTCTGCTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db      304  GCAGGTCAAGGTTTCAAAAGTTAATCAAGAAATGTGATGAACGCTATTCAGAGGAAGCT 363
Qy      448  AGCCAGGCACCAAGTAAAG-----AATGTTGAGGCTACTGTGGCT 486
Db      364  GGAGAACTACCTAATCAGTCGGTGGATAAAACCTCGGAAATGTCGTAGCTCCCTCGCA 423
Qy      487  GCTACTATTGAGCAAGAGGATTTTCACAAATATCGAGNAATACACCAAGTGA---- 541
Db      424  TCTGAGTCTTTGGACAATGTGCTGTTCGCGCTTAAAGATGCTCAACCACTGTCAACCTG 483
Qy      542  -----TAGAACCCCAAGCAAAACAAATACAGGCATGACTACTACAACTGTCACCA 594
Db      484  TCCTATCAAGCATCTGCTGCCAGCCAAATACAGCATGAATTTTACCAAGAGCCAGAG 543
Qy      595  GAAAGTGTACTGCAATATTGTTGTAAGGTTTCCTGCTGATAGTGTAGTCAATGATTTT 654
Db      544  GAGTGTGTGTGACTATATTTCGCAAGGGAATACCAAGCCAAAGATGTTGTTGTGACTTT 603
Qy      655  GGTGAACAGATGTTGAGTGTATCCATTGAAGTTTCTGTTGTAAGAACCAATACCATTTT 714
Db      604  GGTGAACAAATACTTGTGTTAGCATGATGTCGCGGGTGACGAACATTTCTTTCCAG 663
Qy      715  CCCCGTCTGTTTCTAAGATTATCCCTGAGAAATGCAATATCAAGTCTTTATCCACCAAG 774
Db      664  CTTAGGTTGTTGGCAAGATAACACCTGCAAAATGCAATATGAAGTGAATGTCACCAAA 723
Qy      775  GTTGAATAGCCCTTGCAAAAGCTGAGCAGGTGACATGGACACCCCTGGATTTAGTGA 834
Db      724  ATTGAGATCCGCTTGCAAAAGCTGAACCCCTTACATGGACATCTCTTGAATATA---CG 780

```


Db 997 AACAAATCAATTGTCAGAGTCGAATGGGACGGTACTGTGACAAACTGAAAGAGTTGGG 1056

Qy 1108 GCAAAGAGCTAGAGGAGCCCTCATGCTGATGCTGAGCTCAGAGTGGGATCTAA 1167

Db 1057 ACTAAGAAAGTGGAGAGCACTCCACCATGATGGCTCAGAAAGTGGGATTTGA 1116

Qy 1168 AGTT 1171

Db 1117 TCYT 1120

RESULT 7

LOCUS AY150487

DEFINITION Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete cds.

ACCESSION AY150487

VERSION AY150487.1

KEYWORDS GI:23297701

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1108)

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabisidopsis Open Reading Frame (ORF) Clones

Unpublished

2 (bases 1 to 1108)

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Location/Qualifiers

1. 1108

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="U09567"

/ecotype="Columbia"

FEATURES

source

gene

1. 1108

/gene="At4g11260"

CDS

1. 1077

/gene="At4g11260"

/codon_start=1

/evidence=experimental

/product="unknown protein"

/protein_id="AA12904.1"

/db_xref="GI:23297702"

/translation="MAKELAEKAEAFLEDDFDVAVDLYSKAIDLDPNCAFFADRAQ ANIKIDNFTFAVVDANKALEFTLAKAYLRKGTACMKLEYSYAKALEKGSVAPN RPEFKMIDKDLRIAEKEDLVQPMPPSPSSSTTVPLATEADAPPVPIPAAPAKPMF RPEFYQKPEAVVTIYAKVPKENVTVVEGEQILSVVDVAGEBAVHLPRLFKLIP EKRFVLSYKVEIRLAKASIIITWASLEYGKGSVLKPNVSSALSRSRPVYPSKPAK DWKLGKAEVKKQDEKDLGDAAAGFFSDIYSSADEDMRRMNKKSFAESNGTVLSTN WKEVGTGKVESTPPDGMELKKWEY"

1078. 1108

/gene="At4g11260"

3'UTR

1. 1108

/gene="At4g11260"

ORIGIN

Query Match 28.8%; Score 418.4; DB 8; Length 1108;

Best Local Similarity 62.9%; Pred. No. 8.9e-63;

Matches 682; Conservative 0; Mismatches 396; Indels 6; Gaps 2;

Qy 88 GCCTCGGATCTGAGAGCAAGGCCNAGAGCGCTTCGTCGACGACGCTTCGAGCTGGCC 147

Db 4 GCCAAGGAATTAGCAGAGAAGCTTAAGAAGACTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy 148 ACCGAGCTCTACAGCAGGCGCATCGACCGCGGGCCGCCACCGCGGACCTCTATGCCGAC 207

Db 64 GTTGACTTATCTCCAAAGCCATTGACTTGATGCCCAATTCGCCGCTCTTCGCCGAT 123

Qy 208 CGCGCCCGGCGCACATCAAGCTCGGCACTACACTGAGGCTGTGGCGGATGCTTAACAAA 267

Db 124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCCAACAA 183

Qy 268 GCATTTGAGCTTGCATCTATGATGCATAAAGCTTACTACCGGAAGGTGCTGCATGATTT 327

Db 184 GCCATTTGAGTTGAGCGCAACGTTGGCAAAAGCCCTATCTCAGAAAGGCGACTGCTTGTATG 243

Qy 328 AAGCTTGAAGAATACCAAACTGCAAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCAATCA 387

Db 244 AAGCTAGAGAAATATAGTACTGCTTAAGCAGCCCTGGAAGAGGAGCTTCTGTGTGACCG 303

Qy 388 GCGCATTCAGGTTTGTCTCTCTATTGAAGAAATGTGATGAGCGCATCGGTGAGGAATCT 447

Db 304 AATGAACCAAGTTTAAAGAGATGATAGATGGAATGCGATCTTGTATTTGCAGAGAAGAG 363

Qy 448 AGCAGGCGCAGTAAGAATGTTGAGGCTACTGTGCTGCTACTATTGAGGACCAAGAG 507

Db 364 A---AAGATTGTTGTCAGCGGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy 508 GATTTCCAAATATGAGAAATACACCAACAGTATAGAACCCCAAGCAACCAACCAATAC 567

Db 421 GCAACGGAAGCTGATGCTCTCTCTGTTCCAAATTCCTGCAGCACTGCGCAACCAAGTTC 480

Qy 568 AGGCATGACTACTACAAACAGTGCACAGAAAGTGGTACTGCAATATTGCTTAAGGGTGT 627

Db 481 AGGCAGAGTTCTACCCAGAAACAGAAAGCGGTGGTGACAAATTTTCGCAAAAAGTA 540

Qy 628 CTTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687

Db 541 CCTAAGGAGAAACGCTAAGTGTGAGTTTGGTGAAGAGATTCCTGAGTGTGTGCTGATGTT 600

Qy 688 CTTGTTGAGAAACCATACCAATTTTCAGCCCGCTCTGTTTCTTAAGATTATCCCTGAGAA 747

Db 601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCCGGGAAGATAATACCAGAGAG 660

Qy 748 TGCAAAATATCAAGTCTTATCCCAAGGTTGAAATACGCTCTGCAAAAGCTGACGAGGTG 807

Db 661 TGCAGATTGAAGTATTGTGCAACCAAGTTGAGATCCGTCTTTCGGAAGAGCAGATATTC 720


```

Db      958  AACAAATCATTTGCAGAGTCGATGGACGGTACTCTCGACAACTGGAAAGAGTTGGG 1017
QY      1108  GCAAAGAGTAGAAGGAGCCCTGCTGATGGTATGAGTCTCAAGAGTGGGAATCTAA 1167
Db      1018  ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGAGCTCAAGAGTGGGAGTATTGA 1077
QY      1168  AGTT 1171
Db      1078  TCTT 1081

```

```

RESULT 9
AF428340
LOCUS   Arabidopsis thaliana AT4g11260/F8L21_50 mRNA linear PLN 18-OCT-2001
DEFINITION Arabidopsis thaliana AT4g11260/F8L21_50 mRNA, complete cds.
ACCESSION AF428340.1 GI:16226817
VERSION   AF428340.1
KEYWORDS  FLI CDNA.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

```

```

REFERENCE
AUTHORS   Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
          Bowers, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
          Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
          Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
          Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
          Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
          Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
          Davis, R.W., Theologis, A. and Ecker, J.R.
          Arabidopsis cdna clones
          Unpublished
          2 (bases 1 to 1400)

```

```

TITLE     Arabidopsis cdna clones
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 1400)
AUTHORS   Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
          Bowers, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
          Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
          Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
          Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
          Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
          Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
          Davis, R.W., Theologis, A. and Ecker, J.R.
          Direct Submission

```

```

TITLE     Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory
JOURNAL   (SIGAL), Plant Biology Laboratory, The Salk Institute for
          Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
          USA

```

```

COMMENT   RIKEN Genomic Sciences Center (GSC) members carried out the
          collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
          Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
          Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
          Hayashizaki, Y. and Shinozaki, K.

```

```

          The Salk, Stanford, RPEC (SSP) Consortium members carried out the
          sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H.,
          Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowers, L.,
          Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
          Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
          Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
          Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
          Davis, R.W., Theologis, A., and Ecker, J.R.

```

```

          Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
          this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
          contributed equally to this work as PIs.

```

```

FEATURES
source   Location/Qualifiers
1. 1400
    /organism="Arabidopsis thaliana"
    /mol_type="mRNA"
    /db_xref="taxon:3702"
    /chromosome="4"

```

```

/clone="RAFL09-50-H05 (R14746)"
/ecotype="Columbia"
1. .131
132. .1208
/notes="unknown protein"
/codon_start=1
/product="AT4g11260/F8L21_50"
/protein_id="RAL16270.1"
/db_xref="GI:16226818"

```

```

5'UTR
CDS

```

```

/translation="MAKELAEKAEKAEFLDDDDFDVADLYSKAIDLDPNCAAFADRAQ
ANIKIDNFTAEVVDVANKAIELEPTLAKYLRKGTACMKLEEYETAKAALKEKASVAPN
EPKFKMIDECDLRIAEBEKDLVQPMPPSPSSSTPLATEADAPPVPIPAAPAKPMF
RHEFYOKPEAVVTIPAKVPKENTVTFGEQILSVVIDVAGEEAHLQRLFGKIIIP
EKCRPEVLSKVEIRLAKAIIITWASLEYKGSQVLPKNVSNLSALSORPVYPSKPKAK
DWDKLEAEVKKQEKDEKLDGDAAMNKFSDIYSSADEDEMRANWKSFAESNGIVLSTN
WKEVGTGKVKVESTPPDGMELUKWEY"
misc_difference 703
/notes="compared to genome"
/replaces="g"
1209. .1400

```

```

3'UTR
ORIGIN

```

```

Query Match      28.8%; Score 418.4; DB 8; Length 1400;
Best Local Similarity 62.9%; Pred. No. 8.9e-63;
Matches 682; Conservative 0; Mismatches 396; Indels 6; Gaps 2;

```

```

QY      88  GCCTCGGATCTGGAGAGCAAGGCCAAGGAGCCCTTCGTGACAGCACTTCGAGCTGGCC 147
Db      135  GCCAAGAAATAGCAGAGAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 194
QY      148  ACCGAGCTCTACAGCCAGCCATCGACGCGGCGCCGCCACCGCCGACCTCTATGCCGAC 207
Db      195  GTTGACTTATCTCCAAAGCCATTGAGTCCCAATTCGCCCGCCCTTCCTGCCGAT 254
QY      208  CGCGCCAGCGGCACATCAAGCTCGGCACACTACACTGAGGCTGTGGCGGATGCTAAACAA 267
Db      255  CGTGCTCAGGCCCAACATCAAAATCGATACTTCACTGAAGCTCTGTGAGATGCAACAA 314
QY      268  GCATTTGAGCTTGATCCTATGATGCAATAAGCTTACTACCGAAAGGTGCTGCATGATT 327
Db      315  GCCATTTGAGTTGGAGCAACGTTGGCAAGCCATCTCAGAAAGGCGCACTGCTGTGATG 374
QY      328  AGCTTTGAGATACCAAACTCAAGAGGCTGCTCTGAGTTGGGTCTCTTATGATCA 387
Db      375  AGCTAGAGAATATAGTACTGCTAAAGCAGCCCTGGAAGGAGGCTCTCTGTGCACCG 434
QY      388  GCGGATTCAGGTTTCTGCTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db      435  AATGAACCAAGTTTAAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
QY      448  AGCCAGGCAACAGTAAAGAAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db      495  A---AAGATTTGGTTTCAGCGGATGTCACCGAGTTTGCCTTCAAGCTCTACACACCACTA 551
QY      508  GATTTTCAAAATATGGAGATATACACACAGTATAGAACCCCAAGCAACCAACCAATAC 567
Db      552  GCAACCGAAGCTGATGCTCTCTCTGTTTCCAAATTCCTGCGAGCACCCTGCAAAACGATGTC 611
QY      568  AGCGATGACTACTACAAAGTGCACAGAGTGGTACTGACAAATATTGCTAAAGGGTGT 627
Db      612  AGGCAGAGTTCTACAGAAACAGAGAGGCGGTGTCGATGATGATGATGATGATGATGATGAT 671
QY      628  CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Db      672  CTTAAGGAACCTTAACCTGCGAGTTTGGTGGAGCAGATTCGAGTGTGTTGTTGTTGTTGTT 731
QY      688  CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
Db      732  GCTGGAGAGGAGGAGCTTATCATCTCCAGCGGAGATGTTTGGGGAAGATAATACCAGAGAG 791
QY      748  TGCATAATATCAAGTCTTATCCCAAGGTTTGAATACGCTTCCAAAGCTGAGCAGGTG 807
Db      792  TGCAGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851

```



```

QY 808 ACATGACAACTCGATTATAGTGAAGCAAGCAAGTGTCCCGAGAGATAGCAGC 867
Db 852 ACCTGGGCTCCCTTGAATATGTAAGGCAAGTGTTCGCCCAACCCCAATGTCTCA 911
QY 868 CCAGCTGAACACAGCCCAAGACCTTCATACCCATCTTCAAGGCGAAAAAGACTGGGAT 927
Db 912 TCAGC---GCTGTGAGAGACAGTGTACCCATCTTCTAAGCCAGCAAGACTGGGAC 968
QY 928 AACTGGAAGCTGAAGTCAAAAGAGGAGAGAGAAAGAACTTGTATGTGTGCA 987
Db 969 AAGTGTGAGCTGAAGTGAAGAAACAGGAGAGGATGAGAAGCTTGTGAGAGATGCAGCT 1028
QY 988 TTGAACAAATCTTCGTGATCACTACAGGATGCTGATGAAGATATGGGAGGCGCATG 1047
Db 1029 ATGAACAAATTTTTTCAGCGACATATCTCGAGTGCAGATGAAGACATGAGCGGCAATG 1088
QY 1048 ATGAAGTCATCTCTGGATCAATATGACCTGTCTCTCAACCAATTTGGAAGATTTGGA 1107
Db 1089 AACAAATCATTTTCAGAGTCGATGGGACGGTACTGTGACAACTTGGAAAGAGTTGG 1148
QY 1108 GCAAAGAGTAGAAGGAGGACCCCTCGATGATGATGAGCTCAAGAGTGGGATACTAA 1167
Db 1149 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGATGGAGCTCAAGAAGTGGGAGTATTGA 1208
QY 1168 AGTT 1171
Db 1209 TCTT 1212

RESULT 10
AF370229 Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete
LOCUS AF370229 1420 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete
cde.
ACCESSION AF370229
VERSION AF370229.1 GI:13877932
KEYWORDS FLI CNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1420)
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Barth, J., Chung, M.K.,
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 1420)
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Barth, J., Chung, M.K.,
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (18-APR-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X.,
Sakano, H., Pham, P.K., Barth, J., Chung, M.K., Goldsmith, A.D.,

```

Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES

source

1..1420

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="RAF104-14-007 (R09567)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."

1..1420

/gene="At4g11260"

1..136

/gene="At4g11260"

misc_difference 32..34

/gene="At4g11260"

/note="not present in genomic sequence"

137..1213

/gene="At4g11260"

/codon_start=1

/evidence=experimental

/product="unknown protein"

/protein_id="AAK44044.1"

/db_xref="GI:13877933"

/translation="MAKELAKAEKAEFLDDDFDVAVDLYSKAIDLDPNCAAFADRAQ
ANIKDNTFEAVDANKALELEPTLAKAVLRKGTCTMKLEVSTAKALEKASVAPN
EPFKMIDCEDLRIAEKOLVQPMPSLSSSTPLATEADAPVPIPAKPMF
RHEFYQPEEAIVTIPAKVKPVNVEFGEQILSVVIDVAGEEYHLOPRLFGKLIIP
EKCRFEVLSVKVEIRLAKABITWASLEYGKGQSVLPKPNVSSALSQRVYPSKPAK
DWDKLEAEVKQEKDEKLDGDAAANKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
WKEVGTGKVESTPPDGMELKKWEY"

1214..1420

/gene="At4g11260"

ORIGIN

Query Match 28.8%; Score 418.4; DB 8; Length 1420;
Best Local Similarity 62.9%; Pred. No. 8.9e-63;
Matches 682; Conservative 0; Mismatches 396; Indels 6; Gaps 2;

QY 88 GCGTCGATCTGGAGAGCAAGGCCAAGGAGGCTTGTGTCGACGACGCTTCGAGCTGGCC 147

Db 140 GCCAAGGAATTAGCAGAGAAAGCTAAAGAGCTTTTCTAGATGATGACTTCGATGTGCT 199

QY 148 ACCGAGCTCTACAGCCAGCCGATCGACGCGGCGCCGCCACCCGACCTCTATGCCGAC 207

Db 200 GTTGACTTATCTCCAAAGCCATGATCTGATCCCAATTCGCCGCTTCTTCGCCGAT 259

QY 208 CGCGCCGACGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGCGGATGCTAACAA 267

Db 260 CGTGTCTCAGGCCCAACATCAAAATCGATACTTCACTGAAGCTGTCTAGATGCGAACAA 319

QY 268 GCAATGAGCTTGCCTCTATGATGATCAATGCTTACCGGAAAGTGTCTGCATGCATT 327

Db 320 GCCATTTGAGTTGGAGCCAGCTTTGGCAAGCCCTATCTCAGAAAGGCGACTGCTTGTATG 379

QY 328 AAGCTGAAGAATACCAAACTGCAAAAGGCTGCTCTTGGATTGGTTCCTTTATGATCA 387

Db 380 AAGCTAGAGAAATATAGTACTGCTTAAGCAGCCCTGAAAAGGAGGCTTCTGTTGACCG 439

QY 388 GCGGATTCAGGTTTGTCTGCTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGATCT 447

Db 440 AATGAACCAAGTTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499


```
QY 448 AGCCAGCACCAGTAAGATGTTGAGGCTACTGTGCTGCTACTATTGAGGACAAGGAG 507
Db 500 A---AAGATTGGTTGCGCATGCGCATGCCAGGTTGGCTTCAAGCTCTACACACACTA 556
QY 508 GATTTTCAAAATATGGAGATAACACCAAGTAGTAGAAGCCCAAGCAAAACCAAAATAC 567
Db 557 GCAACGGAAGCTGATGCTCTCTGTTTCCAAATTCCTGCAGCACCCTGCCAAACCGATGTC 616
QY 568 AGCGATGACTACTACAAAGTGCACAGAAAGTGGTACTGACATATTTGCTAAGGGTGT 627
Db 617 AGCGACGAGTTCTACCAAGAAACAGAGAAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 676
QY 628 CTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAGTT 687
Db 677 CTTAAGAGAACGTAACGTGCGAGTTTGGTGGAGAGATTCGTAGTGTGTCTATGATGTT 736
QY 688 CTGTGTGAAGAACATACCAATTTTACGCCCGCTGCTGTTTCTAAGATTATCCCTGAGAAA 747
Db 737 CTGTGAGAGAGCTTATCATCTCCAGCCGAGATTGTTCCGGGAAGATATACCAAGAGAG 796
QY 748 TGCAATATCAAGTCTTATCCCAAGGTTGAAATAGCGCTTGCAAAAGCTGAGCAGGTG 807
Db 797 TGCAGATTTTGAAGTATGTCGACCAAGTTGAGATCGTCTTGGAAAGCAGAGATAATC 856
QY 808 ACATGACAACTCTGATATATAGTGGNAGACCAAGACTGTTCCCGAAGATAGCAGC 867
Db 857 ACCTGGGCTCTCTTGAATATGTAAGGGCAAGTGTGTTTGGCCCAACCCAAATGTCTCA 916
QY 868 CCAGCTGMAACACAGCCCAAGACTTTCATACCCATCTTCAAGGCGCAAAAGAGCTGGAT 927
Db 917 TCAGC---GCTGTGCGAGACCAAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGAC 973
QY 928 AAAGTGAAGCTGAAGTCAAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
Db 974 AAGTTGAAGCTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
QY 988 TTGAACAAATCTTCCGTGACATCTACAGAGATGCTGATGAGATATCGGAGGCGCATG 1047
Db 1034 ATGAACAAATTTTTCAGCGACATATCTCGAGTGCAGATGAAGACATGAGCGCGCAATG 1093
QY 1048 ATCAAGTCAATCTGTGAATCAATAGGACATGCTCTCTCAACCAATTTGGAAGATGTTGA 1107
Db 1094 AACAAATCAATTCAGAGTGCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1153
QY 1108 GCAAGAGAGTAGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1167
Db 1154 ACTAAGAACTGGAGAGACTCCACAGATGGCATGGAGCTCAAGAGTGGAGTATTGA 1213
QY 1168 AGTT 1171
Db 1214 TCTT 1217

RESULT 11
AX412470
LOCUS 1077 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 234 from Patent WO022675.
ACCESSION AX412470
VERSION AX412470.1 GI:21444928
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
Glazebrook, J., Wang, X., Dangl, J. L., Eulgem, T. and Zhu, T.
Plant genes, the expression of which are altered by pathogen
infection
Patent: WO 022675-A 234 21-MAR-2002;
SYNGENTA PARTICIPATIONS AG (CH); UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Xun (US); Dangl,
```

```
Jeffrey L. (US); Eulgem, Thomas (US)
Location/Qualifiers
1..1077
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 28.7%; Score 417.6; DB 6; Length 1077;
Best Local Similarity 63.0%; Pred. No. 1.2e-62;
Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

QY 88 GCGTCGATCTCGAGAGCAAGGCAAGGAGGCTTGTGTCAGCAGCAGCTTCGAGCTGCCC 147
Db 4 GCCAAGATATGACAGAGAAAGCTTAAGAGCTTTTCTAGATGATGACTTCGATGTTGCT 63
QY 148 ACCGAGCTCTACAGCAGGCGCATCGAGCGCGCGCCGCCACCGCGCCTCTATATGCCGAC 207
Db 64 GTTGACTTATCTCCAAAGCCATTGACTTGGATCCCAATTTGCGCGCTTCTTCGCGGAT 123
QY 208 CGCGCCAGCGGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db 124 CGTGCTCAGGCCCAACATCAAAATCGATAACTTCACTGAAGCTGTGTAGATGCGAACA 183
QY 268 GCAATTTGAGCTTGATCTCTATGATGATAAAGCTTTACTACCGGAAAGGTGCTGCATGATT 327
Db 184 GCCATTGAGTTGGAGCCAAAGTTGGCAAAAGCCCTATCTCAGAAAGGCGACTGCTTGATG 243
QY 328 AAGCTTTGAAGAATACCAAACTGCAAAAGGCTGCTCTTGGTGGTGTCTTCTTATGCAATCA 387
Db 244 AAGCTAGAGATATAGTACTGCTTAAGCAGCCCTGGAAAGGAGGCTTCTGTTGCAACCG 303
QY 388 GCGATTTCAAGGTTTGTCTGCTTATTTGAAGAAATGATGAGCGCATCGCTGAGGAATCT 447
Db 304 AATGAACCAAAAGTTTAAAGAGATGATAGATGAATGCGATCTTCTGTTTGCAGAAAGAG 363
QY 448 AGCCAGCACCAGTAAAGATGTTGAGGCTACTGTGCTGCTACTATTGAGGCAAGGAG 507
Db 364 A---AAGATTGGTTGAGCGGATGCCACCGAGTTTGCCTTCAAGCTCTACACACCACTA 420
QY 508 GATTTCACAAAATATGAGAGATACACCAAGTAGTAGAAGCCCAAGCAAAACCAAAATAC 567
Db 421 GCRAAGAGCTGATGCTCTCTCTGTTTCCAATTCCTGCAGCACCCTGCCAAACCGATGTC 480
QY 568 AGCATGACTACTACAAAGTGCACAGAAAGTGGTACTGACATATTTGCTAAGGGTGT 627
Db 481 AGCGACGAGTTCTACCAAGAAACAGAAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540
QY 628 CTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAGTT 687
Db 541 CCTAAGGAGAACGTAACGTGCGAGTTTGGTGAAGCAGATTCGTAGTGTGTGCTGATGTT 600
QY 688 CTGTGTGAAGAACATACCAATTTTACGCCCGCTGCTGTTTCTAAGATTATCCCTGAGAAA 747
Db 601 GCTGAGAGAGAGCTTATCATCTCCAGCCGAGATTGTTCCGGGAAGATATACCAAGAGAG 660
QY 748 TGCAAAATATCAAGTCTTATCCCAAGGTTGAAATAGCGCTTGCAAAAGCTGAGCAGGTG 807
Db 661 TGCAGATTTGAAGTATTTGCGACCAAGTTGAGATCCGTCTTTCGAAAGCAGAGATAATC 720
QY 808 ACATGACAAACCTCGGATATATAGTGAAGACCAAGACTGTTCCCGAAGATAGCAGC 867
Db 721 ACCTGGGCTCTCCCTTGAATATGTAAGGGCAAGGTTGTTTGGCCCAACCCAAATGTCTCA 780
QY 868 CCAGCTGAAACACAGCCCAAGACTTTCATACCCATCTTCAAGGCGCAAAAGAGCTGGAT 927
Db 781 TCAGC---GCTGTGCGAGAGACCGATGTTACCCATCTTCTTAAGCCAGCAAAAGACTGGAC 837
QY 928 AAAGTGAAGCTGAAGTCAAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
Db 838 AAGTTGAGAGCTGAAGTGAAGAAACAGGAGAGAGATGAGAGAGCTTGTATGAGATGACGCT 897
QY 988 TTGAACAAATTTTCTCGTGTGATCTATCAAGGATGCTGATGAAGATATGCGAGGCGCATG 1047
```

Db 898 ATGAACAAATTTTTCAGCGCATATCTCGAGTGCAGATGAAGACATGAGCGGCAATG 957
 QY 1048 ATGAAGTCATCTCGTGGATCAATATGGCACTGTTCTCTCAACCAATTCGAAGATGTTGGA 1107
 Db 958 AACAAATCATTTGAGAGTGAATGGAGCGGTACTGTGCAAACTGGAAGAGTTGGG 1017
 QY 1108 GCAAGAGGTAGAGAGGAGCGCCCTCGATGGTATGAGCTCAAGAGTGGGAATCTAA 1167
 Db 1018 ACTAAGAAAGTGGAGAGCACTCCACAGATGGATGGAGCTCAGAGTGGAGATTGA 1077

RESULT 12

AJ620883
 LOCUS 1197 bp mRNA linear PLN 16-JAN-2004
 DEFINITION Brassica oleracea mRNA for SGT1-like protein (sgtlb-2 gene).
 ACCESSION AJ620883
 VERSION AJ620883.1 GI:40974916
 KEYWORDS SGT1-like protein; sgtlb-2 gene.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

1 Yemm A.A., Holub E. and Tor M.

Unpublished

2 (bases 1 to 1197)

REFERENCE

Yemm A.A.
 Direct Submission
 Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research,
 Horticultural Research International, Wellesbourne, Warwick, UNITED
 KINGDOM

FEATURES

source

Location/Qualifiers

1..1197

/organism="Brassica oleracea"

/mol_type="mRNA"

/variety="Senna"

/db_xref="taxon:3712"

/tissue type="leaf"

1..1197

/gene="sgtlb-2"

1..1065

/gene="sgtlb-2"

/function="role in disease resistance"

/codon_start=1

/evidence=experimental

/product="SGT1-like protein"

/protein_id="CAF06581.1"

/db_xref="GI:40974917"

/translations="MASELAEKAEAFDDPDVAADLYSKAIDLPSCAFPADRAQ
 ANIKLNPTEAVADANKATELEPTLAKVLRKACMKLEEVATAKALEKASVAPN
 ESKPEKMDIECNLLIAEERKDLVQVPPPLPSSSTPLALADAPPAAPAKNFRHEP
 YQPEVGVAFKANGPKONVDFGQILSVVDVAGEAVHFQRLFGKILPEKCR
 YEVLTKVIRLAKSIVTWASLEYKGGQALLPKPNVASAVSQRPYPYSSKPKQWDK
 LEAEVKQEKDKLDGDAANKFFSDIYQSADEDMERRAMNKSFAESNGTVLSTNWEV
 GTKVKESTPPDGMELMKWEY"

gene

CDS

ORIGIN

Query Match 28.4%; Score 413.8; DB 8; Length 1197;
 Best Local Similarity 63.6%; Pred. No. 5.6e-62;
 Matches 687; Conservative 1; Mismatches 373; Indels 20; Gaps 3;

QY 88 GCCTCGGATCTCGAGAGCAAGGCAAGCGCTTCGTGACGACGACCTTCGAGCTGGCC 147
 Db 4 GCAGCGAATACGGAAGAAAGCTAAGAGCGCTTCTTAGATGACGACTTCGATGGCT 63
 QY 148 ACCGAGCTCTACAGCGAGGCATCGAGCGCGCGGCCACCGCCGACCTCTATGCCGAC 207
 Db 64 GCTGACTTATCTACCAAGCCATTGACTTGATCCCTCTTTCGCGCTCTCTTCGCGAT 123
 QY 208 CGGCGCCAGCGGACATCAAGCTCGGCACTACACTGAGGCTGTGGGAGTGTACAA 267

Db 124 CGTGCTCAAGCCACACATCAAACTCTTTAACTTCAACGAGCTGTTCAGATGCCAACAA 183
 QY 268 GCATTTGAGCTTGCATCCTATGATGATGAAGCTTACTTACCGGAAGAGTGTGATGCATT 327
 Db 184 GCCATCGAGTTGGAGCTACTTTGGCCAAAGCTTATCTCAGAAAGGCGCTGCTTGTATG 243
 QY 328 AAGCTTTGAAGAAATACCAAACTGCAAAAGGCTGCTCTTGAAGTGGGTTCTTCTTATGCA 387
 Db 244 AAGCTAGAGAGTATGCTACTGCTTAAGCAGCTCTTGAAGAGGCGCTCCGTTGCACCA 303
 QY 388 GGGGATTCAGGTTTGTCTGCTTATTTGAAGAAATGATGAGCGCATCGCTGAGGAATCT 447
 Db 304 AATGAATCCAAAGTTTGAAGAGATCATAGATGAATGCAACCTTCTCATTTGAGAAGAA 360
 QY 448 AGCCAGGCGACCAAGAAATGTTGAGGCTACTGTGCTGCTACTATTGAGGACCAAGGAG 507
 Db 361 -----GAGAAAGATTGGTTCAACAGGTGCCACCGACTTTTGCCTTCAAGCTC 407
 QY 508 GATTTTCAAAATATGGAGAATA-CACCACCAAGTGATAGAACCCCAAGCAAAACCAATA 566
 Db 408 TACAACACCACTAGCAATCGCAGCTGATGCTCTCTCGAGCACCTGSCCAACCCATGTT 467
 QY 567 CAGGCAATGACTACTACAACAGTGCACAGAGTGGTACTGACAAATATTTCCTAAGGGTGT 626
 Db 468 CAGACACGAGTTTCTACCAAGCCGGAAGAGTGGGGTGGCAATTTTCGCTAAAGGAT 527
 QY 627 TCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 686
 Db 528 CCNAAAGCAGACGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587
 QY 687 TCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746
 Db 588 TGCTGAGAGAGAGCTTATCAITTTCCAGCAAGATTGTTCCGGAAGATAATACCAAGAA 647
 QY 747 ATGCAAAATATCAAGTCTTATCCCAAGGTTGAAATAGCGCTTGCAAAAGCTGAGCAGGT 806
 Db 648 GTGCAGATATGAGTATTTGTCGACCAAGTTGAGATCCGCTCTCGCAAAAGCAGAGATAG 707
 QY 807 GACATGCAACACCTCGATTATAGTGAAGACCAAGACCTGTTCCCGCAGAGATAAGCAC 866
 Db 708 CACTTGGGCTCTCCCTTGATA---TGGCAAGGGCAAGCTCTTCTGCTTAACCCATGT 764
 QY 867 GCCAGCTGAAACAGCCCAAGACCTTATACCCATCTTCAAGCGCGGAAAGAAAGCTGGGA 926
 Db 765 TGCCATCAGCAGTGTCTCAGAGGCGGTGTACCCATCTCTTAAGCCGGGAAAGACTGGGA 824
 QY 927 TAACTGGAAGCTGAAGTCAAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 986
 Db 825 CAAAGCTGGAAGCTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
 QY 987 ATTGAACAAATCTTCGCTGACATCTACAAGGATGCTGATGAAGATATCGGAGGGCCAT 1046
 Db 885 TATGAACAAAGTTTTCAGGACATATACAGAGTGCAGATGAGACATGAGGCGAGCCAT 944
 QY 1047 GATGAAGTCAATTCGTGGAATCAAAAGCGCTGTTCTTCTCAACCAATTTGGAAGATGTTG 1106
 Db 945 GAACAAATCATTTGAGAGTCCAAATGGGACGGTGTCTTACAAACTGGAAGAGGTTGG 1004
 QY 1107 AGCAAGAGGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1166
 Db 1005 GACAAAGAGAGTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1064
 QY 1167 A 1167
 Db 1065 A 1065

RESULT 13

AJ620882
 LOCUS 1068 bp mRNA linear PLN 16-JAN-2004
 DEFINITION Brassica oleracea mRNA for SGT1-like protein (sgtlb-1 gene).
 ACCESSION AJ620882
 VERSION AJ620882.1 GI:40974914

KEYWORDS SGT1-like protein; sgt1b-1 gene.
 SOURCE Brassica oleracea
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 Yemm, A.A., Holub, E. and Tor, M.
 AUTHORS Unpublished
 JOURNAL 2 (bases 1 to 1068)
 REFERENCE Yemm, A.A.
 AUTHORS Direct Submission
 TITLE Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research, Horticultural Research International, Wellesbourne, Warwick, UNITED KINGDOM
 FEATURES
 source Location/Qualifiers
 1..1068
 /organism="Brassica oleracea"
 /mol_type="mRNA"
 /variety="Senna"
 /db_xref="taxon:3712"
 /tissue type="leaf"
 1..1068
 /gene="sgt1b-1"
 1..1068
 /gene="sgt1b-1"
 /functions="role in disease resistance"
 /codon_start=1
 /evidence=experimental
 /product="SGT1-like protein"
 /protein_id="CAF06580.1"
 /db_xref="GI:40974915"
 /translation="MASELAEKAEFLDDFDVAVLYSRAIDLPNCAFFADRAQ ANIKLFTSAVDANADIALEPTLSKYLKGTACMKLEYSYAKAALOKGASVAPN ESKFKLIDENLHIAEEKDLAQOMPTLPSSSTTPPLATAADSPAPAKMFERHE FYQPERVTVTFKAGIPKQNLNVEFGDQILSVVIDVAGBEAHFQRLPGKLIIPDK RYEVLSKVEIRLAKBIIITWASLEYVKGQALLPKPNVASVRSOPRYPSSKPAKMD KLEAVKQEKDEKLDGDAANKFFSDIYQSADEMRMANMKFPAESNGTVLSTNWE VGTKKVSTPPDGMELKKWEY"
 ORIGIN
 Query Match 28.0%; Score 407.2; DB 8; Length 1068;
 Best Local Similarity 63.1%; Pred. No. 7.9e-61;
 Matches 682; Conservative 0; Mismatches 383; Indels 15; Gaps 3;
 QY 88 GCGTCGGATCGGAGACGAGCCCAAGGAGCGCTTCGTCGACGACGACTTCGAGCTGCGC 147
 DB 4 GCGAGCGAATGACGCGAAAAGCTTAAGAGAGCTTCTTAGAGGACGACTTCGATGTCGCC 63
 QY 148 ACCGAGCTCTACAGCGAGCCCATCGAGCGCGCGCCGACCGCGGACCTCTATGCGGAC 207
 DB 64 GTTGACTTATCTCCAGAGCCATTGACTTGGATCCCAATTGCGCGCGCTTCTTCGCGGAT 123
 QY 208 CGCCCGCAGCGGACATCAAGCTCGGCACATACACTAGAGGTGTGGCGATGCTAACAAA 267
 DB 124 CGTCTCAGGCGCAACATCAAAATCTTAACTTTACCGAAGCTGTTCGAGATGCGCAAAA 183
 QY 268 GCAATTGAGCTTGATCTCTATGATCATAAAGCTTACTACCGAAGGTGCTGCGATGATT 327
 DB 184 GCCATTGAGTGGAGCTTACGTTGCTTAAGACTTACTCTCAGAAAGGCACTGCTTGATG 243
 QY 328 AAGCTTGAAGAATACCAAACTGCAAGGCTCTCTTGAGTTGGGTCTTCTTATGCGATCA 387
 DB 244 AAGCTAGAGGATACAGTACTGCTTAAGACGACCTTCAAGGCGGCTTCGTTGCGACCA 303
 QY 388 GCGGATCAAGGTTGCTGCTATTTGAAGAAATGTGATGAGCGCATCGCTGAGGAATCT 447
 DB 304 AATGAATCAAGTTTAAACAAAGTTGATAGTGAATGCAATCTTCAATTCGAGAAGAAGAG 363
 QY 448 AGCCAGGACACGATTAAGAAATGTTGAGCTACTGTGGCTGCTACTATTGAGGACAGAG 507
 DB 364 AAGATTGTCGTCACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 417

508 GATTTCACAAATATGGAGAATACACACACAGTATAGAGACCCCAAGCAGACCAAAATAC 567
 DB 418 CCACTAGCAACTGAGCTGATTTCTCTCTCTG-----CAGCACCTGCCAAACCCATGTT 471
 QY 568 AGGCATCACTACTACAAACAGTGCACAGAAAGTGTACTGCAAAATATTTGCTAAAGGGTGT 627
 DB 472 AGACACGAGTTCTACACAGAGCGGAGAAAGTGTGTGTGACAGTTTTCGCTAAAGGGATA 531
 QY 628 CTGCTGATAGTGTAGTCAATTGATTTTGGTGAACAGATGTTGATGTATTCATTGAAGTT 687
 DB 532 CCAAAGCAGAACTTTGAATGTGAGTTTGGTGTATCAGATTTCTGAGTGTGTGATTTGATGT 591
 QY 688 CTGCTGTAAGAACATACCACTTTTACGCCCGCTGCTGTTTCTTAAGATTTATCCCTGAGAAA 747
 DB 592 GCTGGAGAGGAAGCTTATCACTTCCAGCAAGATTGTTGGGAAGATAATACCAAGACAG 651
 QY 748 TGCAATATCAAGTCTTTATCCACCAAGGTTGAATAGCCCTTGCAAAAGCTGAGCAGGTG 807
 DB 652 TGCAGATATGAGTATTTGTCAACCAAGTTGAGATCCGCTCTCGCAAAAGCAGAGATAATC 711
 QY 808 ACATGGACAACCTTGGATTATAGTGGAAAGCCAAAGACTGTTCCTCCAGAGATTAAGCAG 867
 DB 712 ACTTGGGCTCTCTTGAATA---TGTCAAAGGGCAAGCTTCTTCTGCTTAAACCCCAATGTC 768
 QY 868 CCAGCTGAACAGCCCAAGACCTTATACCCATCTTCAAGGCGAAAAGACTGGGAT 927
 DB 769 GCATCAGCGTTTTCGAGAGCGCGGTGTACCCATCATCTAAGCCGCGGAAAGACTGGGAC 828
 QY 928 AACTGGAAGCTGAAGTCAAAAAGGAGGAGAGGAAGAAACCTTGAATGGTGTGATGCTGCA 987
 DB 829 AGCTGGAGGCTGAAGTGAAGAAACAGAGAGAGATGAGAACTAGATGAGATGCGAGCT 888
 QY 988 TTGAACAAATTTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATCGCGAGGCGCATG 1047
 DB 889 ATGAACAAAGTTTTCACGCGATATATACAGAGTCTGATGAGGATATGAGCGGAGCATG 948
 QY 1048 ATGAAGTCACTGTTGGAATCAATGGCAGCTTCTCTCAACCAATGGAAGATGTTGGA 1107
 DB 949 AACAAATCATTTGCTGAGTCAAAATGGACAGTGTCTACGAACTGGAAGAGAGTTGGG 1008
 QY 1108 GCAAGAAGGTAGAGGAGCGCCCTGATGGATGAGCTCAAGAGTGGGAATACTAA 1167
 DB 1009 ACTAAGAAGTGGAGAGCACTCCACAGATGTTATGAGCTCAAGAAAGTGGAGTATTGA 1068
 RESULT 14
 AF439975 1291 bp mRNA linear PLN 06-NOV-2003
 LOCUS Arabidopsis thaliana SGT1a (SGT1a) mRNA, complete cds.
 DEFINITION Arabidopsis thaliana SGT1a (SGT1a) mRNA, complete cds.
 ACCESSION AF439975
 VERSION AF439975.1 GI:17017307
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1291)
 AUTHORS Arevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P.
 TITLE The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance
 JOURNAL Science 295 (5562), 2073-2076 (2002)
 MEDLINE 21893744
 PUBMED 11847307
 REFERENCE 2 (bases 1 to 1291)
 AUTHORS Arevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre, Colney Lane, Norwich NR47UH, UK
 FEATURES
 source Location/Qualifiers
 1..1291

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1. 1291
/gene="SGT1a"
77..1129
/cds="SGT1a"
/codon_start=1
/product="SGT1a"
/protein_id="AAL33611.1"
/db_xref="GI:17017308"
/translation="MAKELADKAKEAFVDDFDVAVDLYSKAIDLDPNCAEFADPRAQ
AYIKLESFTBAVDANAKAIELDPSLTAYIKRGKTACMKLEIRYAKTALKEGASITPS
ESFKPLIDCNFLITEERDLQVPSFTSPVAPPSELDPVTPTAKYRHYQKP
EEVVTTFAGKIDPKQNNVIDGQILSVIEVPGEDAYILQPLRFGLKIIIPDKCKYVL
STKIEIRLADGIIITWASLEHGKPAVLPRKPNVSSEVSORPAYPSSKKVWDKLEAE
VKQEKDEKLEGAALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTWQEVGKTK
IESTPPDGMELKWEI"

```

ORIGIN

```

Query Match      27.1%; Score 394.8; DB 8; Length 1291;
Best Local Similarity 62.8%; Pred. No. 1.1e-58;
Matches 678; Conservative 0; Mismatches 372; Indels 30; Gaps 3;

```

```

Qy 88 CGCTCGGATCTGGAGACGAAGCCCAAGAGGCGCTTCGTGACGACGACTTCGAGCTGGCC 147
Db 80 GCGAAGGAGCTTGTCTGATAGGCTTAAAGAGCGCTTCGTAGACGATGACTTCGATGTGCT 139
Qy 148 ACCGAGCTCTACAGCCAGGCGCATCGAGCGCGGCGCCGCCACCGCGACCTTATGCCGAC 207
Db 140 GTTGACTTGTACTCCAAGCCATGACTTGGATCCTAAATTCGCGTGAGTCTTCGCTGAT 199
Qy 208 CGCGCCGAGCGGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTTAACAAA 267
Db 200 CGTGCTCAGCGCTATATCAACTCGAAGCTTCACTGAGCGCGTGGCAGATGCCAACAAA 259
Qy 268 GCAATTGAGCTTGATCTCTATGATGATCAATAAGCTTACTACCGGAAAGTGTGCTGATGATT 327
Db 260 GCAATTGAGTTGGATCTCTCATTTGACCAAAAGCTTACCTAAGAAAAGAACTGCGCTGTATG 319
Qy 328 AAGCTTCAAGAAATACCAAACTGCAAGGCTGCTCTGAGTTGGTCTTCTTATGATCA 387
Db 320 AAGCTTGAAGATATCGAGCTGCTTAAACAGCTCTTGAAGGGTGTCTTATCACCGCG 379
Qy 388 GCGGATTCAGAGTTGCTGCTCTATTTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db 380 AGTGAATCCAAATTTAAGAAGTTGATAGATGAATCAATTTCTTAATCACAGAAGAA--- 436
Qy 448 AGCCAGCAGCAGTAAGAATTTGAGGCTACTGTGCTGCTACTATTGAGGACAAGGAG 507
Db 437 -----GAGAAAGATTTGGTTCAACCGGTGCTTCGACTTTGCTTCAAGTGT 483
Qy 508 GATTTCAAAATATGGAGATATACACCAAGTGTATGAACCCCAAGCAACCAAAATAC 567
Db 484 GACAGCACCA-----CCAGTATCTGAACCTTGATGTACCCCTACAGCCAAAGTAC 532
Qy 568 AGGCATGACTACTACACAGTGCACACAGTGTGTCTGACAAATTTTGTCTAAGGGTGT 627
Db 533 AGGCACGAGTACTACCAAGGCGAGAGTGTGTGTAAGTGTCTTTTGAAGGAATA 592
Qy 628 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Db 593 CCCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652
Qy 688 CTTGTTGAAGAACATACCAATTTTTCAGCCCGCTGCTGTTTCTAAGATTTCCCTTGAGAA 747
Db 653 CTTGAGAGAGATGCTGATTTATCTCAACCGAGATGTTTGAAGATATATCCAGATTAAG 712
Qy 748 TGCAAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCTTGCAAAAGCTGAGCAGGTG 807
Db 713 TGCAAAATATCAAGTCTTATGTCGACCAAAATTTGAGATCGTCTTGCAAAAGCGATATATC 772
Qy 808 ACATGGCAACCCCTGGATTTATAGTGGAGAACCAAGAACTGTTCCCGAAGATTAAGCAG 867

```

```

Db 773 ACATGGGCTCCCTCGAACACGCGCAAGGCGCGGTTTTCGCAAGCCCTAATGTCTCA 832
Qy 868 CCAGCTGTAAGACAGCCCAAGACCTTTCATACCCATCTTCAAAGGCGCAAAAAGACTGGAT 927
Db 833 TCAGAGGTTTCA---CAGAGGCGAGCTTATCCATCTTCTAAGAAAGTTAAGGACTGGAC 889
Qy 928 AAATCGAAGCTGAAGTCAAAAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 987
Db 890 AAGCTGGAAGCTGAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCT 949
Qy 988 TTGAACAAATCTTCCTGATCATCTACAGGATGCTGATGAAGATATCGGAGGCGGCGCATG 1047
Db 950 TTGAACAAATCTTCCTGAGATATACAGAAATCGGATGAGGATATGAGGCGGAGCGATG 1009
Qy 1048 ATGAAGTCATTCTGGAATCAAAATGGCAGCTTCTTCAACCAATTCGAAAGATGTTGGA 1107
Db 1010 AGCAATCATTTGGATCGAATGGACAGTCTGTCAACAAATGCGAAGAGTTGGG 1069
Qy 1108 GCAAGAGAGTAGAAGGAGGCGCCCTGATGATGAGCTCAAGAGTGGGAATATAA 1167
Db 1070 ACTAAGACAATCGAGAGTACTCTCCGATGGCATGGAGCTCAAGAAATGGAGATCTGA 1129

RESULT 15
LOCUS AY096395
DEFINITION Arabidopsis thaliana putative phosphatase (At4g23570) mRNA, complete cds.
ACCESSION AY096395
VERSION AY096395.1 GI:20465860
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1084)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1084)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGECC (SFP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

```

Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PCBC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PCBC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

Location/Qualifiers

1..1084

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="U11621"

/ecotype="Columbia"

/note="This clone is in pUNI 51."

1..1084

/gene="At4g23570"

1..1053

/gene="At4g23570"

/codon_start=1

/evidence=experimental

/product="putative phosphatase"

/protein_id="AA020035.1"

/db_xref="GI:20465861"

/translation="MAKELADAKAEAFVDDDFVAVDLYSKAIDLDNCAEPFADRAQ

AYKLESFTAVADANAIELDPISLRKAYLRKGTACMKLEBYRTAKTALSKGASITPS

ESFKFLIDCNFLITEEKDLVQVPSTPLSPVAPPSELDVPTPKRHYHYQKP

EEVVTVFAKADIPKQNIIDFGEQILSVIWEVGEDAYLQPLRFGKIIIPDKCKEVL

STKIEICLADIIITWASLEHSGKPAVLKPNVSSEVSQRPAPSSKKVMDWDLKLE

VKQEKDEKLEADALNKFREIYQNADEDMRRMSKSFVESNGTVLSTWQEVGYKT

IESTPPDGMELKWEI"

1054..1084

/gene="At4g23570"

ORIGIN

Query Match 27.0%; Score 393.2; DB 8; Length 1084;

Best Local Similarity 62.7%; Pred. No. 2.1e-58;

Matches 677; Conservative 0; Mismatches 373; Indels 30; Gaps 3;

```
QY 88 GCGTCGGATCTGGAGAGCAAGGCCAAGCGCCCTTCGTGACGACGAGCTTCGAGCTGGCC 147
DB 4 GCGNAGGAGCTTGCTGATAGGCTTAAGAGCTTCGTAGCAGATGACTTCGATGTTGCT 63

QY 148 ACCGAGCTCTACAGCCAGGCCATCAGCGCCGGCCCGCCACCGCCGACCTCTATGCCGAC 207
DB 64 GTTGACTTGACTCCAAAGCCATTGACTTGGATCCTAATTGGCGCTGAGTCTTCGCTGAT 123

QY 208 CGGCCAGGCGGACATCAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
DB 124 CGTCTCAGGCTATATCAAACTCGAAAGCTTCACTGAGGCCGTGGCAGATGCGAACAAA 183

QY 268 GCAATTGAGCTTGATCTCTATGATCATAAAGCTTACTACCGAAAGGTGCTGCAATGATT 327
DB 184 GCAATTGAGTTGATCTCTTATTGACCAAGCTTACTACGAAAGGAACTGCTCTGATG 243

QY 328 AAGCTTGAAGAAATACCAAACTGCAAGGCTGCTCTTGGATTTGGTTCTTCTTATGATCA 387
DB 244 AAGCTTGAAGAGTATCGGACTGCTTAAACAGCTCTTGAAGAGGTGCTCTATCAACGCG 303

QY 388 GGCATTCAGGTTTGCTGCTTATTGAGGAATGTGATAGCGCATCGCTGAGGAATCT 447
DB 304 AGTGAATCCAAATTAAGAAAGTTGATAGTAATGCAATTTTCTAATCACAAGAA-- 360

QY 448 AGCCAGGACCAAGTAAGAAATGTTGAGGCTACTGTGCTGCTACTATTGAGGCAAGGAG 507
DB 361 -----GAGAAAGATTTGGTTCAACCGGTGCTTCGACTTTGCGCTTCAAGTGT 407

QY 508 GATTTCACAAATATGGAGAAATACACACCGAGTATAGAACCCCAAGCAAAACAAATAC 567
DB 408 GACAGCACCA-----CCAGTATCTGAATCTGATGTGTCACCCCTACAGCCAAAGTAC 456
```

```
QY 568 AGCATTGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTGCTAAGGGTGT 627
DB 457 AGGACGAGTACTACCAAAAGCCAGAGAGTGTGGTAACTGTTTTGCAAAAGGAATA 516

QY 628 CTTGCTGATAGTGTAGTCAATTGATTTTGGTGAACAGATGTTGAGTGTATCAATTGAAATT 687
DB 517 CCCAAGCAGAAATGTTAACTCGACTTTGGTGAACAAATCTGAGTGTGTTGATGAGGTT 576

QY 688 CTTGGTGAAGAACATACCAATTTTACGCCGCTGTTTCTTCTAGATTTATCCCTGAGAA 747
DB 577 CTTGGAGAGGATCGGTATTATCTCCAAACCGAGATTGTTTGGAAAGATAATACCAGATAAG 636

QY 748 TGCATAATATCAAGTCTTATCCACCAAGGTTGAAATACGCTTGCACCAAGCTGACAGGTG 807
DB 637 TGCATAATATGAGTATTGTCGACCAAAATTTGAGATCTGTCTTGCACCAAGCCGATATATC 696

QY 808 ACATGGACAAACCTGGATTATAGTGAAGACCAAAAGACTGTTCCCAAGAGATTAAGCAG 867
DB 697 ACATGGGCTCTCTCGAAACACGCGCAAGGGCCAGCGGTTTGCACCAAGCCTAATGTCTCA 756

QY 868 CAGCTGAAACAGCCCCCAAGACCTTCATACCCATCTTCAAGGGCAAAAGACTGGGAT 927
DB 757 TCAGAGGTTTCA---CAGAGGCCAGCTTATCCATCTTCTAAGAAAGTTAAGGACTGGAC 813

QY 928 AAATGGAAGCTGAAGTCAAAAAGGAGGAGAGGAAGAAACCTTGATGTTGATGCTGCA 987
DB 814 AAGCTGGAAGCTGAAGTGAAGAGAGCAAGCAAGAAAGATGGAAGCTGGAAGGAGAGCAGCT 873

QY 988 TTGAACAAATTTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGCCATG 1047
DB 874 TTGAACAAATTTCTTCCGGGAGATATACCAGATGCGGATGAGGATATGAGGCGAGCGATG 933

QY 1048 ATGAAGTCAATCTGTTGGAATCAAAATGGCAGCTGTTCTCTCAACCAATTCGAAAGATGTTGA 1107
DB 934 AGCAAAATCATTTGTGGAATCGAATGGGACAGTGTCTGTCACAACTGGCAAGAGTTGGG 993

QY 1108 GCAAGAGAGCTAGAAGGGAGCCCGCTGATGGTATGGAGCTCAAGAAAGTGGGAATACTAA 1167
DB 994 ACTAAGCAATCGAGAGTACTCTCTCCGATGGCATGGAGCTCAAGAAATGGGAGATCTGA 1053
```

Search completed: April 23, 2005, 08:51:26

Job time : 6464.31 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Ruh on: April 22, 2005, 23:55:26 ; Search time 628.264 Seconds
(without alignments)
10232.696 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggcgcgctcgatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N Geneseq_16Dec04:*
2: Geneseqn1980s:*
3: Geneseqn1990s:*
4: Geneseqn2000s:*
5: Geneseqn2001as:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	733.6	67.6	1104	12	ADQ37150 Cell prol
2	733.6	67.6	1104	12	ADQ15734 Rice stre
3	490	45.1	571	10	ADFS1379 Maize rar
4	419.2	38.6	1366	3	AAC39274 Arabidops
5	418.8	38.6	1324	3	AAC40498 Arabidops
6	417.6	38.5	1077	6	ADG87792 A. thalia
7	417.6	38.5	1361	3	AAC51074 Arabidops
8	393.2	36.2	1117	3	AAC42413 Arabidops
9	286	26.3	287	6	ABL71197 Cotton gy
10	231.4	21.3	581	9	ACL25190 DNA clone
11	221.8	20.4	585	13	ACN47711 Cotton pr
12	217.8	20.1	534	13	ACN59636 Cotton gy
13	216.4	19.9	272	6	ABL75330 Corn tass
14	205.6	18.9	541	9	ACL25189 DNA clone
15	194.8	17.9	528	13	ACN59570 Cotton gy
16	194.4	16.9	573	13	ACN47765 Cotton pr
17	183.6	16.9	573	13	ACN47765 Cotton pr
18	180.8	16.6	500	9	ACL25185 DNA clone
19	175.4	16.2	523	13	ACL61180 Cotton gy
20	164.4	15.1	272	10	ABX86113 Corn ear-

21	156.4	14.4	1025	3	AAZ52451	Aaz52451 HTRM clon
22	156.4	14.4	1185	6	ABL42163	AbL42163 Nucleotid
23	156.4	14.4	1264	6	ABQ55038	Abq55038 Human ova
24	156.4	14.4	1366	8	ABZ24700	Abz24700 Human cel
25	156.4	14.4	1535	8	ACD13363	AcD13363 Human DNA
26	156.4	14.4	1555	10	ADD29650	Add29650 Human tum
27	156.4	14.4	1776	5	ABV28161	Abv28161 Human pro
28	156.4	14.4	1776	5	ABV22337	Abv22337 Human pro
29	156.4	14.4	2053	2	AAZ27345	Aaz27345 Human sec
30	156.4	14.4	2053	9	ADA07224	Ada07224 Human cDN
31	156.4	14.4	2053	12	ADN40960	Adn40960 Novel hum
32	148.6	13.7	291	6	ABL73761	AbL73761 Corn tass
33	141.4	13.0	409	8	ABX62506	Abx62506 Arabidops
34	141.4	12.3	615	13	ADQ55550	Adq55550 Novel can
35	129.2	11.9	481	10	AD82158	Ade82158 Arabidops
36	125	11.5	388	10	ADFS1380	Adf1380 Maize rar
37	111.4	10.3	464	10	ADB50274	Adb50274 Primary r
38	102.4	9.4	340	9	ACL25187	AcL25187 DNA clone
39	102.4	9.4	412	9	ACL25184	AcL25184 DNA clone
40	93.2	8.6	672	12	ADJ40154	Adj40154 Plant cDN
41	90.8	8.4	1242	6	ABZ31779	Abz31779 Candida a
42	78.8	7.3	431	6	ABL62696	AbL62696 Colon ade
43	78	7.2	1303	12	ADJ43595	Adj43595 Plant cDN
44	75.8	7.0	518	6	ABQ55648	Abq55648 Human ova
45	75.8	7.0	869	4	AAK63551	Aak63551 Human imm

ALIGNMENTS

RESULT 1
ADQ37150
ID ADQ37150 standard; DNA; 1104 BP.
XX
AC ADQ37150;
XX
DT 07-OCT-2004 (first entry)
XX
DE Cell proliferation-related nucleic acid sequence #155.
XX
KW cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response; ds.
XX
OS Oryza sativa.
XX
FN WO2004061122-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041200.
XX
PR 26-DEC-2002; 2002US-0436565P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
DR WPI; 2004-534388/51.
XX
PT New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
producing enhanced food crops.
XX
PS Claim 57; SEQ ID NO 309; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body

CC	of the specification as EPO data.									
XX										
SQ	Sequence	1104 BP;	323 A;	244 C;	284 G;	253 T;	0 U;	0 Other;		
	Query Match	67.6%;	Score	733.6;	DB	12;	Length	1104;		
	Best Local Similarity	80.9%;	Pred. No.	6e-180;						
	Matches	888;	Conservative	0;	Mismatches	189;	Indels	21;	Gaps	2;
QY	4	GCCCGCTCGGATCTGGAGAGCAAGGCCAAGAGGCGCTTCTGACAGCACTTCGAGCTG	63							
DB	13	GCCCGCTCGGATCTGGAGAGCAAGGCCAAGAGGCGCTTCTGACAGCACTTCGAGCTC	72							
QY	64	GCCACGAGCTCTACAGCGAGGCGATCGAGCGCGGCCCGCCACCGCCAGCTCTATGCC	123							
DB	73	GCCCGCGAGCTCTACAGCGAGGCGATCGAGCGCGGCCCGCCACCGCGAGCTCTAGCC	132							
QY	124	GACCGCGCCAGCGCGCATCAAGCTCGGCAATACATACACTAGGCTGTAGCTGTATGCTAAC	183							
DB	133	GACCGCGCCAGCGCGCATCAAGCTAGGCAATACATACACTAGGCTGTAGCTGTATGCTAAC	192							
QY	184	AAAGCAATTGAGCTTATCTATGATGATCAATAGCTTACTACCGGAAGGTGCTGCATGC	243							
DB	193	AAGGCCATTGAACTTACCCCAATCAATGCAAGGCTTATCTTCTAAGGCGCTGCAATGT	252							
QY	244	ATTAAGCTTGAAGATACCAACTGCAAGGCTCTCTTGAGTTGGTTCTTCTTAAGCA	303							
DB	253	ATACGACTGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATTTGGTTACTGCTTCGCA	312							
QY	304	TCAGGCGATCAAGTTTGTCTTATTAAGGAATGTGATGAGCCATCGCTGAGGAA	363							
DB	313	TCTGTGACTCAAGTTTACTCGCTAATGAAGAGTGTGATGAGCGCATTTGCTGAGGAG	372							
QY	364	TCTAGCGAGGACCAAGTAAGATTTGAGG-----CTACTGTGGCTGCT	408							
DB	373	CTTACTGAGTCCCTGTTAAGAGGCTGAAATGAGGAGCTGCGCCCTCTGTGCTTCT	432							
QY	409	ACTATTGAGGCAAGGAGGATTTCAAAATATGAGGAATACACCAAGTATGATAGRACC	468							
DB	433	TTTGTGAGGAAAGGATGATGCTGCAAAATGATGATATACACCAATGGTAGA----	488							
QY	469	CCAAGCAACCAAAATACAGGCACTACTACACAGTGCACAGAGTGTGTTACTGACA	528							
DB	489	--AGTGAAGCCAAATATACAGGCACTTCTACACAGTGTCTACAGAAATGTTATGACA	546							
QY	529	ATATTGCTGAAGGTGTTCTCTGATAGTGTAGTCAATTTTGGTGAAACAGATGTG	588							
DB	547	ATTTTGAAGGTTGTTCTGCTGAGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	606							
QY	589	AGTGTATCCATTGAAGTTCCTGCTGAGAAACCATACCAATTTTCAAGCCGCTGTTTCT	648							
DB	607	AGTGTGCTGAATGAAGTTCCTGCTGAGAGGAGCGGTACCAATTTTCAAGCCGCTGTTTCT	666							
QY	649	AAGATTTATCCCTGAGAAATGCAAAATATCAAGTCTTATCCCAAGGTTGAAATAGCCTT	708							
DB	667	AAGATCATCCCTGAGAAAGCAGATACCAAGTGTCTATCCACCAAGGTTGAAATAGACTG	726							
QY	709	GCAAAAGCTGAGCAGGTGATGAGCAACCTCTGATTTATAGTGAAGACCAAGACTGTT	768							
DB	727	GCTAAGCTGAACAGATTATCATGGACCTCACTTGTATGATATAAAACCAAGGCTGTT	786							
QY	769	CCCAGAAAGATAGCAGCGGCTGAGAAACAGCCCAAGACCTTCTATCCCATCTTCAAG	828							
DB	787	CCACAAAGATATATCCCTCCAGCTGAATCGGGCCAGAGGCCATCATATCTTCTCAAAA	846							
QY	829	GCGAAAAAGACTGGGATTAACCTGGAAGCTGAAGTCAAAAGGAGGAGGAGGAGAA	888							
DB	847	TCCAAAGAAAGACTGGGATTAACCTGGAAGCTGAAGTCAAAAGGAGGAGGAGGAGAG	906							
QY	889	CTTGATGGTATGCTGATTTGAACAAATTTCTTCCTGATCATCTACAGGATCTGATGAA	948							
DB	907	CTTGAGGCGATGCTGATTTGAACAAATTTTTCCTGATCATCTACAGTATGCTGATGAA	966							
QY	949	GATATGCGAGGCGGCATGATGAAGTCAATCTGTTGAATCAATGGCACTGTTCTCTCAACC	1008							

967 GACATGCGACGAGCAATGATGAAATCTTTTGTGTAATCTTAACGGTACTGTTCTGTCGACC 1026

1009 AATTGGAAGATGTTGGAGCAAGAGGTAGAGGAGCCCTGATGTTGATGAGCTC 1068

1027 AATTGGAAGATGTTGGCTGCAAGAGGTAGAGGAGCCCTGATGTTGATGAGCTT 1086

1069 AAGAGTGGGAATCTAA 1086

1087 AAGAAATGGAGTACTAA 1104

RESULT 2

ADQ15734

ID ADQ15734 standard; DNA; 1104 BP.

XX

AC ADQ15734;

XX

DT 07-OCT-2004 (first entry)

XX

DE Rice stress-related protein coding sequence #72.

XX

XX rice; stress-related protein; plant maturation; plant development;

KW plant proliferation; plant senescence; plant disease-resistance;

KW plant stress response; transgenic plant; pest tolerance;

KW herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;

KW improved nutritional value; increased yield; increased proliferation;

gene; ds.

XX

OS Oryza sativa.

XX

XX WO2004061080-A2.

FN

XX 22-JUL-2004.

PD

XX 23-DEC-2003; 2003WO-US041098.

XX

XX 26-DEC-2002; 2002US-0436564P.

PR

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

PA

XX Cooper B;

PI

XX WPI; 2004-534374/51.

XX

DR P-PSDB; ADQ15735.

XX

XX New isolated nucleic acids and proteins, useful for producing transgenic

PT plants having improved properties, e.g. tolerance to pests, herbicides,

PT or biotic or abiotic stresses, improved nutritional value, or increased

PT yield or proliferation.

XX

PS Claim 45; SEQ ID NO 143; 551pp; English.

XX

XX The invention comprises the amino acid and coding sequences of rice

CC stress-related proteins. The DNA and protein sequences of the invention

CC are useful for regulating and controlling plant maturation and

CC development, including proliferation, senescence, disease-resistance, or

CC stress response. They are also useful for producing transgenic plants

CC having improved properties, e.g. tolerance to pests, herbicides, or

CC biotic or abiotic stresses, improved nutritional value, increased yield

CC or proliferation, or improved structure causing less loss from lodging or

CC shattering. The present DNA sequence encodes a rice stress-related

CC protein of the invention.

XX

SQ Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

Query Match 67.6%; Score 733.6; DB 12; Length 1104;

Best Local Similarity 80.9%; Pred. No. 6e-180;

Matches 888; Conservative 0; Mismatches 189; Indels 21; Gaps 2;

QY 4 GCCCGCTCGGATCTGGAGAGCAAGGCCAAGAGGCGCTTCTGACAGCACTTCGAGCTG 63

DB 13 GCCCGCTCGGATCTGGAGAGCAAGGCCAAGAGGCGCTTCTGACAGCACTTCGAGCTC 72

QY 64 GCCACGAGCTCTACAGCGAGGCGATCGAGCGCGGCCCGCCACCGCCAGCTCTATGCC 123

DB 73 GCCCGCGAGCTCTACAGCGAGGCGATCGAGCGCGGCCCGCCACCGCGAGCTCTAGCC 132

QY 124 GACCGCGCCAGCGCGCATCAAGCTCGGCAATACATACACTAGGCTGTAGCTGTATGCTAAC 183

DB 133 GACCGCGCCAGCGCGCATCAAGCTAGGCAATACATACACTAGGCTGTAGCTGTATGCTAAC 192

QY 184 AAAGCAATTGAGCTTATCTATGATGATCAATAGCTTACTACCGGAAGGTGCTGCATGC 243

DB 193 AAGGCCATTGAACTTACCCCAATCAATGCAAGGCTTATCTTCTAAGGCGCTGCAATGT 252

QY 244 ATTAAGCTTGAAGATACCAACTGCAAGGCTCTCTTGAGTTGGTTCTTCTTAAGCA 303

DB 253 ATACACTGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATTTGGTTACTGCTTCGCA 312

QY 304 TCAGGCGATCAAGTTTGTCTTATTAAGGAATGTGATGAGCCATCGCTGAGGAA 363

DB 313 TCTGTGACTCAAGTTTACTCGCTAATGAAGAGTGTGATGAGCGCATTTGCTGAGGAG 372

QY 364 TCTAGCGAGGACCAAGTAAGATTTGAGG-----CTACTGTGGCTGCT 408

DB 373 CTTACTGAGTCCCTGTTAAGAGGCTGAAATGAGGAGCTGCGCCCTCTGTGCTTCT 432

QY 409 ACTATTGAGGCAAGGAGGATTTCAAAATATGAGGAATACACCAAGTATGATAGRACC 468

DB 433 TTTGTTGAGGAAAGGATGATGCTGCAAAATGATGATATACACCAATGGTAGA---- 488

QY 469 CCAAGCAACCAAAATACAGGCACTACTACACAGTGCACAGAGTGTGTTACTGACA 528

DB 489 --AGTGAAGCCAAATATACAGGCACTTCTACACAGTGTCTACAGAAATGTTATGACA 546

QY 529 ATATTGCTGAAGGTGTTCTCTGATAGTGTAGTCAATTTTGGTGAAACAGATGTG 588

DB 547 ATTTTGAAGGTTGTTCTGCTGAGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 606

QY 589 AGTGTATCCATTGAAGTTCCTGCTGAGAAACCATACCAATTTTCAAGCCGCTGTTTCT 648

DB 607 AGTGTGCTGAATGAAGTTCCTGCTGAGAGGAGCGGTACCAATTTTCAAGCCGCTGTTTCT 666

QY 649 AAGATTTATCCCTGAGAAATGCAAAATATCAAGTCTTATCCCAAGGTTGAAATAGCCTT 708

DB 667 AAGATCATCCCTGAGAAAGCAGATACCAAGTGTCTATCCACCAAGGTTGAAATAGACTG 726

QY 709 GCAAAAGCTGAGCAGGTGATGAGCAACCTCTGATTTATAGTGAAGACCAAGACTGTT 768

DB 727 GCTAAGCTGAACAGATTATCATGGACCTCACTTGTATGATATAAAACCAAGGCTGTT 786

QY 769 CCCAGAAAGATAGCAGCGGCTGAGAAACAGCCCAAGACCTTCTATCCCATCTTCAAG 828

DB 787 CCACAAAGATATATCCCTCCAGCTGAATCGGGCCAGAGGCCATCATATCTTCTCAAAA 846

QY 829 GCGAAAAAGACTGGGATTAACCTGGAAGCTGAAGTCAAAAGGAGGAGGAGGAGAA 888

DB 847 TCCAAAGAAAGACTGGGATTAACCTGGAAGCTGAAGTCAAAAGGAGGAGGAGGAGAG 906

QY 889 CTTGATGGTATGCTGATTTGAACAAATTTCTTCCTGATCATCTACAGGATCTGATGAA 948

DB 907 CTTGAGGCGATGCTGATTTGAACAAATTTTTCCTGATCATCTACAGTATGCTGATGAA 966

QY 949 GATATGCGAGGCGGCATGATGAAGTCAATCTGTTGAATCAATGGCACTGTTCTCTCAACC 1008

Qy	64	GCACCGAGCTCTTACAGCGAGGCGCATTCGACGCGGGGCGCGCACCGCGGACCTCTATGCC	123
Db	73	GCGCCGAGCTCTACAGCGAGGCAATCGAGGCCACCGCGCGGAGCTCTACGCC	132
Qy	124	GACCGCGCCAGGCGCACATCAGCTCGGCNACTACACTGAGGCTGTGGCGGATGCTAAC	193
Db	133	GACCGCGCCAGGCGCCATATCAAGCTAGGCAACTACACTGAGGCTGTAGTGAATGCTAAC	192
Qy	184	AAAGCAATTGAGCTTGATCTCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC	243
Db	193	AAGGCCATTGACTTGGACCCATCAATGCACAAGGCTTATCTTGGTAAGGCGCTGCATGT	252
Qy	244	ATTAAAGCTTGAAGAATACCAAATGCAAAAGCTGCTCTTGAATGAGGTTCTTCTTTATGCA	303
Db	253	ATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATGGGTTACTCGTTCGCA	312
Qy	304	TCAGGCGATTCAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATGCTGAGGAA	363
Db	313	TCCTGGTCACTCAAGGTTTACTTCGCTTAATGAAGGAGTGTGATGAGCGCATTGTGAGGAG	372
Qy	364	TCTAGCCAGGCGCCAGTAAAGATGTTTGAGG-----CTACTGTGGCTGCT	408
Db	373	CTTACTGAATCCCTGTTAAGAGGCTGAAGATGAGCAGTGCCTCTGTGTCTTCT	432
Qy	409	ACTATTGAGGACRAGGAGGATTTTCACAATATGGAGAATACACACAGATGATAGACCC	468
Db	433	TTTCTGTAGGAAAGGATGATGCTGCAAAATGATTAATACACCAATGGTAGA---488	488
Qy	469	CCAAGCAAAACCAAAATACAGGCATGACTCTACAAAGTGCACAGAGTGGTACTGACA	528
Db	489	--AGTGAAGCCAAATACAGCACGACTTCTACAAAGTGTACAGAAAGTTGATTGACA	546
Qy	529	ATATTTCTCAAGGCTGTTCTCTGCTGATGATGTAGTCAATTCATTTGCTGGAACAGATGTC	588
Db	547	ATTTTTGCAAGGCTGTTCTCTGCTGAGAAATGTTGTTGATTTGCTGAACAAAGTTA	606
Qy	589	AGTGTATCCATGGAAGTTCTCTGCTGGAAGAACCATACCATTTTCAGCCCCCTCTGTTTTCT	648
Db	607	AGTGTGTCGATTGAAGTCCCTGGAGAGGCGGTACCATTTTCAGCCTCGTCTGTTTTCT	666
Qy	649	AAGATTATCCTGTGAAATGCAAAATATCAAGTCTTATCCCAAGGTTGAAATACGCGCTT	708
Db	667	AAGATCATCCTGTGAAAGCAGATACCAAGTGTCTATCCCAAGGTTGAAATAAGACTG	726
Qy	709	GCAAAAGCTGAGCAGGTGACATGACCAACCTTGATTTATAGTGAAGACCAAGACTGTT	768
Db	727	GCTAAAGCTGAACAGATTACATGACCTCACTTGATTTATGATAAAACCAAAAGGCTGTT	786
Qy	769	CCCAGAAAGATTAAGCACCGCAGCTGAAACAGCCCCAAGACTTCATACCCCATCTTCAAAG	828
Db	787	CCACAAAGATATCCCTCCAGCTGAATCGGCCAGAGGCCATCATATCCTTCTCAAAA	846
Qy	829	CGCAAAAAGACTGGGTAAACTCGGAAGCTGAAGTCAAAAAGGAGGAGAGGAAGAAAA	888
Db	847	TCCAAGAAAGACTGGGTAAACTCGGAAGCTGAAGTAAAAAGGAGGAGAGGAGGAAG	906
Qy	889	CTTGATGCTGATGCTGCATTGAAACAAATTTCTTCGTCACATCTACAGGATGCTGATGAA	948
Db	907	CTTGAAGCGCATGCTGCATTGAAACAAATTTTTCGTCACATCTACAGTGAATGCTGATGAA	966
Qy	949	GATATCGGAGGGCCATGATGAAGTCAATCGTGAATCAAAATGCACTGTTTCTCTCAACC	1008
Db	967	GACATCGACGAGCAATGATGAATCTTTTGTGAACTCAACGGTACTGTTCTGTGCGACC	1026
Qy	1009	AATTGGAAGAATGTTGGAGCAAAAGAGGTAGAAAGGAGCCCCCTGATGTTGATGGAGCTC	1068
Db	1027	AATTGGAAGAATGTTGGCTCGAAGAAAGGTAGAGGGAGCCCACTGATGGATGGAGCTT	1086
Qy	1069	AAGAAAGTGGGAATACTAA1086	
Db	1087	AAGAAATGGGAGTACTAA1104	

RESULT 3	
ID	ADFS1379 standard; cDNA; 571 BP.
XX	
AC	ADFS1379;
XX	
DT	12-FEB-2004 (first entry)
XX	
DE	Maize Rar interactor 2 (RarInt2) DNA seq id 2.
XX	
KW	plant protectant; gene therapy; Rarl; disease resistance; plant;
XX	immunogen; maize; Rar interactor 2; RarInt2; gene; ss.
OS	Zea mays.
XX	
FN	US2003167504-A1.
PD	04-SEP-2003.
XX	
PF	12-FEB-2002; 2002US-00074473.
XX	
PR	12-FEB-2001; 2001US-0268157P.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Crane EH;
XX	
DR	WPI; 2003-898130/82.
XX	
PT	New maize Rarl-interactor nucleic acid useful for enhancing disease
PT	resistance in plants, as a probe or amplification primer, in recombinant
PT	expression of polypeptides, or as an immunogen in the preparation or
PT	screening of antibodies.
XX	
PS	Claim 1; SEQ ID NO 2; 30pp; English.
XX	
CC	The invention describes an isolated nucleic acid capable of interacting
CC	with Rarl. The nucleic acid comprises a member selected from: (a) a
CC	polynucleotide comprising a sequence of 388 (S1), 571 (S2) or 691 (S3) bp
CC	fully defined in the specification; (b) a polynucleotide amplified from a
CC	zea mays nucleic acid library using primers that selectively hybridise,
CC	under high stringency conditions, to loci within the polynucleotide in
CC	(a); (c) a polynucleotide encoding a maize Rarl-interactor protein; (d) a
CC	polynucleotide having at least 80 or 90% sequence identity to (a); (e) a
CC	polynucleotide which hybridises under high stringency conditions to (a);
CC	and (f) a polynucleotide complementary to any of the above
CC	polynucleotides. The nucleic acid molecule and methods are useful in
CC	enhancing disease resistance in plants. The nucleic acid molecule may
CC	also be used as a probe or an amplification primer, in recombinant
CC	expression of polypeptides, or as an immunogen in the preparation and/or
CC	screening of antibodies. This sequence represents a maize Rar interactor
CC	2 (RarInt2) polynucleotide.
XX	
SQ	Sequence 571 BP; 151 A; 154 C; 161 G; 105 T; 0 U; 0 Other;
	Query Match 45.1%; Score 490; DB 10; Length 571;
	Best Local Similarity 100.0%; Pred.No. 8.1e-117;
	Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGCGCCGGTTCGAATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTGCGACGACTTCGAG 60
DB	82 ATGCGCCGGTTCGAATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTGCGACGACTTCGAG 141
QY	61 CTGGCCACCGAGCTCTACAGCAGGCCCATCGACCGGGCCCGCACCGCGGACCTCTAT 120
DB	142 CTGGCCACCGAGCTCTACAGCAGGCCCATCGACCGGGCCCGCACCGCGGACCTCTAT 201
QY	121 GCCGACCGCGCCGAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCCGAGTGCT 180
DB	202 GCCGACCGCGCCGAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCCGAGTGCT 261
QY	181 AACAAAGCAATTGAGCTTGATTCCTTATGATGCTATAAGACTTTACTACCGGNAAGGTGCTGCA 240


```
Db      954  ATGAACAAATTTTTCAGCGCATATCTCGAGTGCAGATGAAGACATGAGCGGGCAATG 1013
Qy      967  ATGAAGTCATTTCGTGGAAATCAAAATGGCACTGTTCTCTCAACCAATTCGAAAGATGTTTGA 1026
Db      1014  AACAAATCATTTCGAGAGTCGAATGGACGGTACTGTCGAAACTGGAAGAAGTTGGG 1073
Qy      1027  GCAAAGAGGTAGAGGAGGCCCTCTGATGGTATGGAGCTCAAGAAAGTGGGAATACAA 1086
Db      1074  ACTAAGAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAAGTGGAGTATTGA 1133

RESULT 5
AAC40498
ID  AAC40498 standard; DNA; 1324 BP.
XX
AC  AAC40498;
XX
DT  17-OCT-2000 (first entry)
DE  Arabidopsis thaliana DNA fragment SEQ ID NO: 28522.
XX
KW  Hybridisation assay; genetic mapping; gene expression control;
KW  protein identification; signal transduction pathway; metabolic pathway;
KW  promoter; termination sequence; ss.
XX
OS  Arabidopsis thaliana.
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
XX  25-FEB-2000; 2000EP-00301439.
XX
PR  25-FEB-1999; 99US-0121825P.
PR  05-MAR-1999; 99US-0123180P.
PR  09-MAR-1999; 99US-0123548P.
PR  23-MAR-1999; 99US-0125788P.
PR  25-MAR-1999; 99US-0126264P.
PR  29-MAR-1999; 99US-0126785P.
PR  01-APR-1999; 99US-0127462P.
PR  06-APR-1999; 99US-0128234P.
PR  08-APR-1999; 99US-0128714P.
PR  16-APR-1999; 99US-0129845P.
PR  19-APR-1999; 99US-0130077P.
PR  21-APR-1999; 99US-0130449P.
PR  23-APR-1999; 99US-0130510P.
PR  23-APR-1999; 99US-0130891P.
PR  28-APR-1999; 99US-0131449P.
PR  30-APR-1999; 99US-0132048P.
PR  30-APR-1999; 99US-0132407P.
PR  04-MAY-1999; 99US-0132484P.
PR  05-MAY-1999; 99US-0132485P.
PR  06-MAY-1999; 99US-0132486P.
PR  06-MAY-1999; 99US-0132487P.
PR  07-MAY-1999; 99US-0132863P.
PR  11-MAY-1999; 99US-0134256P.
PR  14-MAY-1999; 99US-0134218P.
PR  14-MAY-1999; 99US-0134219P.
PR  14-MAY-1999; 99US-0134221P.
PR  14-MAY-1999; 99US-0134370P.
PR  18-MAY-1999; 99US-0134768P.
PR  19-MAY-1999; 99US-0134941P.
PR  20-MAY-1999; 99US-0135124P.
PR  21-MAY-1999; 99US-0135353P.
PR  24-MAY-1999; 99US-0135629P.
PR  25-MAY-1999; 99US-0136021P.
PR  27-MAY-1999; 99US-0136332P.
PR  28-MAY-1999; 99US-0136782P.
PR  01-JUN-1999; 99US-0137222P.
PR  03-JUN-1999; 99US-0137528P.
PR  04-JUN-1999; 99US-0137502P.
PR  07-JUN-1999; 99US-0137724P.
PR  08-JUN-1999; 99US-0138094P.
PR  10-JUN-1999; 99US-0138540P.
PR  10-JUN-1999; 99US-0138847P.
PR  14-JUN-1999; 99US-0139119P.
PR  16-JUN-1999; 99US-0139452P.
PR  16-JUN-1999; 99US-0139453P.
PR  17-JUN-1999; 99US-0139492P.
PR  18-JUN-1999; 99US-0139454P.
PR  18-JUN-1999; 99US-0139455P.
PR  18-JUN-1999; 99US-0139456P.
PR  18-JUN-1999; 99US-0139457P.
PR  18-JUN-1999; 99US-0139458P.
PR  18-JUN-1999; 99US-0139459P.
PR  18-JUN-1999; 99US-0139460P.
PR  18-JUN-1999; 99US-0139461P.
PR  18-JUN-1999; 99US-0139462P.
PR  18-JUN-1999; 99US-0139463P.
PR  18-JUN-1999; 99US-0139750P.
PR  21-JUN-1999; 99US-0139763P.
PR  21-JUN-1999; 99US-0139817P.
PR  22-JUN-1999; 99US-0139899P.
PR  23-JUN-1999; 99US-0140353P.
PR  23-JUN-1999; 99US-0140354P.
PR  24-JUN-1999; 99US-0140895P.
PR  28-JUN-1999; 99US-0140823P.
PR  29-JUN-1999; 99US-0140991P.
PR  30-JUN-1999; 99US-0141287P.
PR  01-JUL-1999; 99US-0141842P.
PR  01-JUL-1999; 99US-0142154P.
PR  02-JUL-1999; 99US-0142055P.
PR  08-JUL-1999; 99US-0142390P.
PR  08-JUL-1999; 99US-0142803P.
PR  09-JUL-1999; 99US-0142920P.
PR  12-JUL-1999; 99US-0142977P.
PR  13-JUL-1999; 99US-0143542P.
PR  14-JUL-1999; 99US-0143624P.
PR  15-JUL-1999; 99US-0144005P.
PR  16-JUL-1999; 99US-0144085P.
PR  16-JUL-1999; 99US-0144086P.
PR  19-JUL-1999; 99US-0144325P.
PR  19-JUL-1999; 99US-0144331P.
PR  19-JUL-1999; 99US-0144332P.
PR  19-JUL-1999; 99US-0144333P.
PR  19-JUL-1999; 99US-0144334P.
PR  20-JUL-1999; 99US-0144352P.
PR  20-JUL-1999; 99US-0144632P.
PR  20-JUL-1999; 99US-0144884P.
PR  21-JUL-1999; 99US-0144814P.
PR  21-JUL-1999; 99US-0145086P.
PR  21-JUL-1999; 99US-0145088P.
PR  22-JUL-1999; 99US-0145085P.
PR  22-JUL-1999; 99US-0145087P.
PR  22-JUL-1999; 99US-0145089P.
PR  22-JUL-1999; 99US-0145192P.
PR  22-JUL-1999; 99US-0145145P.
PR  23-JUL-1999; 99US-0145218P.
PR  23-JUL-1999; 99US-0145224P.
PR  26-JUL-1999; 99US-0145276P.
PR  27-JUL-1999; 99US-0145913P.
PR  27-JUL-1999; 99US-0145918P.
PR  27-JUL-1999; 99US-0145919P.
PR  28-JUL-1999; 99US-0145951P.
PR  02-AUG-1999; 99US-0146386P.
PR  02-AUG-1999; 99US-0146388P.
PR  02-AUG-1999; 99US-0146389P.
PR  03-AUG-1999; 99US-0147038P.
PR  04-AUG-1999; 99US-0147204P.
PR  04-AUG-1999; 99US-0147302P.
PR  05-AUG-1999; 99US-0147192P.
PR  05-AUG-1999; 99US-0147260P.
PR  06-AUG-1999; 99US-0147303P.
PR  06-AUG-1999; 99US-0147416P.
```


XX (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
XX Lalgudi RV, Ito LY, Sherman BK;
XX WPI; 2002-163647/21.
XX
XX Novel purified corn tassel-derived polynucleotide useful for determining
PT altered gene expression, to recover regulatory elements and to follow
PT inheritance of desirable characteristics through hybrid breeding
PT programs.
XX
XX Claim 1; SEQ ID NO 571; 201pp; English.
XX
XX The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful in
CC the evaluation, and alteration of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality and yield, and as molecular markers for studying inheritance of
CC multigene traits in a plant breeding program. (I) can be used to produce
CC a tassel-specific profile of gene transcription, a transcript image, to
CC clone regulatory elements for use in transformation vectors, to express a
CC polypeptide, to identify, isolate or extend identical or related corn
CC tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences
XX
XX SQ Sequence 287 BP; 89 A; 60 C; 70 G; 68 T; 0 U; 0 Other;
SQ
Query Match 26.3%; Score 286; DB 6; Length 287;
Best Local Similarity 100.0%; Pred. No. 5.7e-64;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
216 AGCTTACTACCGGAAGGCTGTCATGTCATTAAGCTTGAAGATACCAAACTGCAAGGC 275
Db 1 AGCTTACTACCGGAAGGCTGTCATGTCATTAAGCTTGAAGATACCAAACTGCAAGGC 60
276 TGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTCTGCTCTATTGAA 335
Db 61 TGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTCTGCTCTATTGAA 120
336 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCCACGTAAGAAGATGTTGAGGC 395
Db 121 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCCACGTAAGAAGATGTTGAGGC 180
396 TACTGTGGCTGCTACTATTGAGGCAAGGAGATTTTCAAAATATGAGATACACACC 455
Db 181 TACTGTGGCTGCTACTATTGAGGCAAGGAGATTTTCAAAATATGAGATACACACC 240
456 AGTGATAGACCCCAAGCAAAACCAAAATACAGGCATGACTACTAC 501
Db 241 AGTGATAGACCCCAAGCAAAACCAAAATACAGGCATGACTACTAC 286
RESULT 10
ACL25190/c
ID ACL25190 standard; DNA; 581 BP.
XX
XX ACL25190;
AC
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX

DE DNA clone originating in barley containing SNP encoding sequence #15181.
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Haruna Nijo).
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 27-SEP-2002; 2002JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYN1-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fip.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX SQ Sequence 581 BP; 157 A; 151 C; 101 G; 172 T; 0 U; 0 Other;
SQ
Query Match 21.3%; Score 231.4; DB 9; Length 581;
Best Local Similarity 82.6%; Pred. No. 1.2e-49;
Matches 265; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
766 GTTCCCCAGAGATAAGCAGCCGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCA 825
Db 569 GCTCCTCAGAGATAAATGTACCAGCTGAATCAGCCAGAGGCCATCTTATCTTCATCA 510
826 AAGCGGAAAAAAGACTGGGATAACTCGAAGCTGAAGTAAAAAGGAGGAGAGAGAA 885
Db 509 AAATCCAAAAAGGACTGGGATAAGCTTGAAGCTTGAAGTAAAAAAGAGAGATGAG 450
886 AAACCTTGATGCTGATGCTGCATTGAAACAAATTCCTCGTGCACATCTACAAGGATGCTGAT 945
Db 449 AAACCTTGACGCTGATGCTGCATTGAAACAAATTCCTCGTGCACATCTACAAGGATGCTGAT 390
946 GAAGATATGCGGAGGGCCATGATGAAGTCAATTCGTGGAATCAATGGCAGCTGTTCTCTCA 1005
Db 389 GAAGATATGCTAGAGCAATGATGAAGTCTTTGTGGAGTCTAATGGAACCGTCTCTCA 330
1006 ACCAATTGAAAGATGTTGAGCAAGAGGTGAAGAGGGAGCCCCCTGATGTTGATGAG 1065
Db 329 ACCAATCTGGAAGATGTCGGGAAAAAAGACGCTTGAAGGAAAGCCCTCTGTGTAATGAG 270
1066 CTCAGAGAGTGGGAATACTAA 1086

269 CTCAGAAGTGGGAGTATTAA 249

seqdata.uspto.gov/sequence.html?DocID=US20040123340

RESULT 11

ACN47711
ID ACN47711 standard; cDNA; 585 BP.

ACN47711;
02-DEC-2004 (first entry)

Cotton primed seed EST Clone ID: LIB3825-014-Q1-K6-D9, SEQ.2492.

Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
variety DP50B; library LIB3825; molecular tag; molecular marker;
genetic mapping; molecular mapping; seed germination; plant growth;
plant quality; plant yield; plant breeding; tissue printing; ss.

Gossypium hirsutum.

US2004123340-A1.

24-JUN-2004.

12-DEC-2001; 2001US-00021323.

14-DEC-2000; 2000US-0255619P.

(DEIK/) DEIKMAN J.

(FENG/) FENG P C C.

(FINC/) FINCHER K L.

(ZIEG/) ZIEGLER T E.

Deikman J, Feng PCC, Fincher KL, Ziegler TE;

WPI; 2004-479808/45.

New isolated nucleic acid molecule that encodes a plant protein or its
fragment, useful for isolating a variety of agronomically significant
genes associated with plant growth, quality or yield, and as molecular
tags to map genes.

Claim 1; SEQ ID NO 2492; 34pp; English.

The invention relates to 17880 cotton expressed sequence tags (ESTs;
ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
from primed or non-primed seeds from variety DP50B, mature seeds from
variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
tissue, developing fibres, carpel walls and septa from variety
Nucotton33B. The invention also relates to substantially purified
proteins or their fragments encoded by nucleic acid molecules of the
invention, and to transformed plants having a nucleic acid construct
comprising a nucleic acid of the invention. The cotton ESTs are useful as
molecular tags to isolate genetic regions, to isolate genes, to map
genes, to determine gene function and to determining whether genes are
members of a particular gene family. The nucleic acid molecules may be
used for isolating a variety of agronomically significant genes
associated with plant growth, quality, yield, and could also serve as
links in metabolic and catabolic pathways. The nucleic acid molecules are
also useful for identifying genes important in initiating and maintaining
seed germination or that may be used to mitigate stresses encountered
during seed germination. The ESTs additionally enable the acquisition of
promoters and cis-regulatory elements which will be useful to express
agronomically significant genes in these tissues and/or other tissues,
and also permits the acquisition of molecular markers useful in breeding
schemes, genetic and molecular mapping, and in cloning of agronomically
significant genes. The nucleic acid molecules are further useful for
detecting the expression level or pattern of a protein or mRNA and for
detecting the presence or quantity of a protein by tissue printing. The
present sequence represents a specifically claimed EST isolated from a
cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
data for this patent did not form part of the printed specification, but
was obtained in electronic format directly from the US patent office at

CC seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX Sequence 585 BP; 186 A; 117 C; 134 G; 148 T; 0 U; 0 Other;
SQ

Query Match 20.4%; Score 221.8; DB 13; Length 585;
Best Local Similarity 62.7%; Pred. No. 3.6e-47; Indels 3; Gaps 1;
Matches 362; Conservative 0; Mismatches 212;

QY 148 CTCGGCAACTACACTGAGGCTGTGGCGGATGTAACAAAGCAATTGAGCTTCCTATG 207

DB 1 CTCACAATCTCCCTGACGCTGTGGCAGAAGCTACAAAGCAATTGAGTTGATCGGTCC 60

QY 208 ATGCATAAGCTTACTACCGAAAGGTGCTGATGCAATTAGCTTGAAGATACCAACT 267

DB 61 ATGTCTAAATCCTACTTGGTAAAGCTATCTGGCTGTATGAAGCTTGGAGGATCAAACT 120

QY 268 GCAAGGCTGCTCTTGAAGTGGTCTTCTTATGATCAGGGGATTCAGGTTTGTCTGT 327

DB 121 GCTAAGGCTGCGTTGGAGACTGGGGCTGCTTTGGCCACCAAGACTCGAGATTTTCCAG 180

QY 328 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCCACAGTAAAGAT 387

DB 181 TTGATTAAAGAAATGTGAGAGCGCAATGTCAGAGGAAATGGTGAGTTACCAAGGAGAGC 240

QY 388 GTTGAGGCTACTGTGGCTCTACTATTGAGGACAAGGAGGATTCACAAATATGAGAT 447

DB 241 TTGGAAGAGGTGCCAACAAATGTTTACCTCGGAAGACATCTCTTCTGCCAGGATATA 300

QY 448 ACACCACCAAGTGA---TAGAACCCCAAGCAACCAAAATACAGGCATGACTACTACAA 504

DB 301 CTGTATCCCATGACTGTTCAGACACCAACCAATCACTTACAGGCATGAATTTATCAG 360

QY 505 AGTGCCACAGAAAGTGTGCTACTGCAATATTGTTGTAAGGGTGTCTCTGCTGATAGTGATC 564

DB 361 AAACACAGAGGAAGTGTGCTGCTCAATATTGTCAGAGGAATACCAAGGAGTGTGTAAA 420

QY 565 ATTGATTTTGGTGAACAGATGTTGAGTGATCCATTGAAGTTTCTGGTGAAGAACCATAC 624

DB 421 GTTAAATATGGTGAACAAATATCTAAAGTGTGCTGTCATGCCCCGGAAGATGCTTAT 480

QY 625 CATTTTCAGCCCCGCTCTGTTTCTAAGATTTATCCCTGAGAAATGCAAAATATCAAGTCTTA 684

DB 481 CATCTCCACCTGCTTATTGGAAGATAATACCTGACAGTGACAGATATGATGTTTG 540

QY 685 TCACCAAGGTGAAATAGCCCTTGCAAAAGCTGAGC 721

DB 541 TCAACCAAAATTGAAATTAGGTTAGCAAAAGGCTGAAC 577

RESULT 12

ACN59636

ID ACN59636 standard; cDNA; 534 BP.

XX ACN59636;

XX AC

XX 02-DEC-2004 (first entry)

Cotton gynoecium tissue EST Clone ID: LIB3829-033-Q1-K6-G10, SEQ.14417.

Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
variety Nucotton33B; library LIB3829; molecular tag; molecular marker;
genetic mapping; molecular mapping; seed germination; plant growth;
plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

CC	hybridisation or amplification technologies, as query sequences to	XX	The present invention relates to oligonucleotide clones originating in
CC	determine homology of known sequences, as probe for use in Southern or	CC	barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC	Northern hybridisation, and to identify the presence of and/or to	CC	(SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC	determine the degree of similarity between two (or more) nucleic acid	CC	varieties, identification of particular varieties and genotype-phenotype
CC	sequences	CC	analysis, isolation of specific genes and creation of new varieties by
XX	Sequence 272 BP; 95 A; 54 C; 70 G; 49 T; 0 U; 4 Other;	CC	transformation of barley varieties with them and production of new barley
XX	Query Match 19.9%; Score 216.4; DB 6; Length 272;	CC	varieties with desired properties. The present sequence represents an
XX	Best Local Similarity 90.9%; Pred. No. 6.6e-46;	CC	oligonucleotide clone DNA sequence featured in the specification. The
XX	Matches 241; Conservative 0; Mismatches 19; Indels 5; Gaps 1;	CC	sequence data for this patent did not form part of the printed
QY	703 CGCCTTGCAAAAGCTGAGAGGTGACATGACACACCTCGATTATATAGTGAAGACCAAG 762	CC	specification, but was obtained in electronic format directly from WIPO
DB	1 CGCCTTGCAAAAGCTGAGAGGTGACATGACACACCTCGATTATATAGTGAAGACCAAG 60	CC	at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
QY	763 ACTGTTCCCGAAGATAAGCAGCCAGCTGAAACAGCCGCCAAGACCTTCATACCCATCT 822	XX	standardise OS field)
DB	61 GCTATTCCCGAAGATAAGCAGCCAGCNGAA-----CACANGACCTTCATACCCATCT 115	XX	Sequence 541 BP; 151 A; 144 C; 89 G; 157 T; 0 U; 0 Other;
QY	823 TCAAGGCGAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGGAA 882	Query Match 18.9%; Score 205.6; DB 9; Length 541;	
DB	116 TCAAGTCAAAAAGGACTGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGGAA 175	Best Local Similarity 84.1%; Pred. No. 5.7e-43;	
QY	883 GAAAACTTGATGCTGATGCTGCAATGCAACAAATCTTCGTTGACATCTACAAGGATGCT 942	Matches 232; Conservative 0; Mismatches 44; Indels 0; Gaps 0;	
DB	176 GAAAACTTGATGCTGATGCTGCAATGCAACAAATCTTCGTTGACATCTACAAGGATGCT 235	QY 811 TCATACCCATCTTCAAGGCGAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAG 870	
QY	943 GATGAAGATATGCGGAGGCGCCATGA 967	DB 539 TCTTATCTTCATCAAAATCCAAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAGAA 480	
DB	236 GATGAAGATATGCGGAGGCGCCCTGGA 260	QY 871 GAGGAGAGGAGAAAACTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 930	
RESULT 14		DB 479 CAGGAGAGGATGAGAACTTGACGGTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 420	
ACL25189/c		QY 931 TACAAGGATGCTGATGAAGATATGCGGAGGCGCATGATGAAGTCAATTCGTTGAAATCAAT 990	
ID ACL25189 standard; DNA; 541 BP.		DB 419 TACAGTGATGCTGATGAAGATATGCGGTAGAGCAATGATGAAGTCTTTTGTGGAGTCTAAT 360	
XX ACL25189;		QY 991 GGCACCTGTTCTCTCAACCAATTTGGAAGATGTTGGAGCAAGAGTGAAGGGAGCCCC 1050	
XX 27-OCT-2003 (revised)		DB 359 GGAACCGTTCTCTCAACCAACTGGAAGATGTTGGAGCAAGAGTGAAGGGAGCCCC 1086	
XX 17-OCT-2003 (first entry)		QY 1051 CCTGATGCTGATGCTCAAGAGTGGGAATACTAA 1086	
DNA clone originating in barley containing SNP encoding sequence #15180.		DB 299 CCTGATGCTGATGCTCAAGAGTGGGAATACTAA 264	
Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;		RESULT 15	
Gene; ss.		ACL25186/c	
Hordeum vulgare; var. (cul.Haruna Nijo).		ID ACL25186 standard; DNA; 571 BP.	
WO2003057877-A1.		XX AC ACL25186;	
17-JUL-2003.		XX 27-OCT-2003 (revised)	
16-DEC-2002; 2002WO-IB005403.		XX 17-OCT-2003 (first entry)	
20-DEC-2001; 2001JP-00387059.		DNA clone originating in barley containing SNP encoding sequence #15177.	
20-DEC-2001; 2001JP-00387131.		Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;	
20-DEC-2001; 2001JP-00403299.		Gene; ss.	
20-DEC-2001; 2001JP-00403300.		Hordeum vulgare; var. (cul.Haruna Nijo).	
27-SEP-2002; 2002JP-00327515.		WO2003057877-A1.	
(UYN1-) UNIV JAPAN OKAYAMA.		17-JUL-2003.	
Sato K, Takeda K, Kohara Y;		16-DEC-2002; 2002WO-IB005403.	
WPI; 2003-587127/55.		20-DEC-2001; 2001JP-00387059.	
Single nucleotide polymorphism sites in barley varieties and DNA		20-DEC-2001; 2001JP-00387131.	
sequences containing them for analysis and identification of barley		20-DEC-2001; 2001JP-00403299.	
varieties and production of barley transformants with desired		20-DEC-2001; 2001JP-00403300.	
characteristics.		27-SEP-2002; 2002JP-00327515.	
Disclosure; SEQ ID XX; 284pp; Japanese.		(UYN1-) UNIV JAPAN OKAYAMA.	
		Sato K, Takeda K, Kohara Y;	
		WPI; 2003-587127/55.	
		Single nucleotide polymorphism sites in barley varieties and DNA	
		sequences containing them for analysis and identification of barley	
		varieties and production of barley transformants with desired	
		characteristics.	

This Page Blank (w/ptc)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:46:01 ; Search time 194.463 Seconds
(without alignments)
9137.976 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggcgctcggtatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	164.4	15.1	272	4	US-09-313-294A-4573
2	156.4	14.4	2053	3	US-09-227-357-45
3	76.2	7.0	198	4	US-09-248-796A-2330
4	74.2	6.8	1735	4	US-09-949-016-1499
5	74.2	6.8	1982	4	US-09-016-434-1067
6	74.2	6.8	1982	4	US-09-825-497A-40
7	74.2	6.8	1987	4	US-09-825-497A-39
8	74.2	6.8	1987	4	US-09-517-779-1
9	73.2	6.7	425	4	US-09-513-999C-4027
10	63.4	5.8	47493	4	US-09-949-016-13241
11	62.2	5.7	989	4	US-09-270-767-12741
12	61.2	5.6	2735	4	US-09-551-974A-101
13	61.2	5.6	2735	4	US-09-565-501A-101
14	61.2	5.6	2735	4	US-09-639-206A-101
15	61.2	5.6	2735	4	US-09-874-923-101
16	61.2	5.6	3012	4	US-09-551-974A-94
17	61.2	5.6	3012	4	US-09-565-501A-94
18	61.2	5.6	3012	4	US-09-639-206A-94
19	61.2	5.6	3012	4	US-09-874-923-94
20	61.2	5.6	3134	2	US-08-533-669A-1
21	61.2	5.6	3134	3	US-08-183-861-1
22	61.2	5.6	3134	3	US-08-022-765-1
23	61.2	5.6	3134	4	US-09-551-974A-1
24	61.2	5.6	3134	4	US-09-565-501A-1
25	61.2	5.6	3134	4	US-09-639-206A-1
26	61.2	5.6	3134	4	US-09-874-923-1
27	61.2	5.6	3134	4	US-08-798-841-1

28 61.2 5.6 4233 4 US-09-551-974A-99 Sequence 99, Appl
29 61.2 5.6 4233 4 US-09-565-501A-99 Sequence 99, Appl
30 61.2 5.6 4233 4 US-09-639-206A-99 Sequence 99, Appl
31 61.2 5.6 4233 4 US-09-874-923-99 Sequence 100, Appl
32 61.2 5.6 4917 4 US-09-551-974A-100 Sequence 100, Appl
33 61.2 5.6 4917 4 US-09-565-501A-100 Sequence 100, Appl
34 61.2 5.6 4917 4 US-09-639-206A-100 Sequence 100, Appl
35 61.2 5.6 4917 4 US-09-874-923-100 Sequence 100, Appl
36 61.2 5.6 4929 4 US-09-551-974A-98 Sequence 98, Appl
37 61.2 5.6 4929 4 US-09-565-501A-98 Sequence 98, Appl
38 61.2 5.6 4929 4 US-09-639-206A-98 Sequence 98, Appl
39 61.2 5.6 4929 4 US-09-874-923-98 Sequence 98, Appl
c 40 58 5.3 7218 1 US-08-232-463-14 Sequence 14, Appl
41 56.4 5.2 2221 3 US-09-301-978C-1 Sequence 1, Appl
42 56.4 5.2 2229 4 US-09-949-016-1647 Sequence 1647, Ap
c 43 56.4 5.2 2274 4 US-09-220-132-188 Sequence 188, App
44 52.6 4.8 2111 4 US-09-949-016-1548 Sequence 1548, Ap
45 51.8 4.8 405 4 US-09-513-999C-11529 Sequence 11529, A

ALIGNMENTS

RESULT 1

US-09-313-294A-4573

; Sequence 4573, Application US/09313294A

; Patent No. 6476212

; GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath V.

; APPLICANT: Ito, Laura Y.

; APPLICANT: Sherman, Bradley K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US

; CURRENT APPLICATION NUMBER: US/09/313,294A

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 7600

; SOFTWARE: PERL Program

; SEQ ID NO 4573

; LENGTH: 272

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6476212 700348654H1

; NAME/KEY: unsure

; LOCATION: 92, 100, 103, 105, 112, 114, 140, 204, 210, 244-245, 250, 257, 261-262,

; OTHER INFORMATION: a, t, c, g, or other

US-09-313-294A-4573

Query Match 15.1%; Score 164.4; DB 4; Length 272;

Best Local Similarity 77.3%; Pred. No. 8.7e-35;

Matches 214; Conservative 0; Mismatches 58; Indels 5; Gaps 2;

QY 326 GTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCAGCAGCAGTAAGA 385

Db 1 GTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCAGCAGCAGTAAGA 60

QY 386 ATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAAGGAGGATTTCAAAATATGGAGA 445

Db 61 ATGTCGAGGCTCTCTGTAGCTGCTACTCTGTGANGACAAGGNGTNTCGCAANTTGGATA 120

QY 446 ATACACACAGTGTAGTAGACCCCAAGCAACCAAAATACAGCATGACTTACTACACA 505

Db 121 ATACACCGCCAGTGGTAGANCCCAAGCAACCAAAATATAGGCATGACTTACTACACA 180

QY 506 GTGCCACAGAAGTGTACTGACAAATTTTCTAGGGTGTTCCTGCTGATAGTGTAGTCA 565

Db 181 GTGCCACAGAATGG---TCCTACATNTTCCNAGGGTGTTCGCTGATAGTGA--TC 235

QY 566 TTGATTTTGTGAACAGATGTTGAGTGTATCCATTGA 602

Db 236 ATTGATTCNNTGATNAGATGTNGNCCCTTCA 272

; EARLIER APPLICATION NUMBER: 60/058,660
 ; EARLIER FILING DATE: 1997-09-12
 ; EARLIER APPLICATION NUMBER: 60/058,661
 ; EARLIER FILING DATE: 1997-09-12
 ; NUMBER OF SEQ ID NOS: 672
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 45
 ; LENGTH: 2053
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-227-357-45

Query Match 14.4%; Score 156.4; DB 3; Length 2053;
 Best Local Similarity 56.3%; Pred. No. 3,7e-32;
 Matches 339; Conservative 0; Mismatches 251; Indels 12; Gaps 2;

QY	479	CAAATAACAGGCATGCTACTACAACAGTGCACAGAGTGGTACTGACAAATATTGCTTA 538
DB	1416	CAAAATCAAGTAGTACTGGTATCAACACAGATCTCAAGTAGTCAATTACACTTATGATCA 1475
QY	539	AGGGTCTTCTGCTGATAGTAGTCAATGATTTTGGTGAACAGATGTTGAGTGTATCCA 598
DB	1476	AGAATGTTCAGAGAAGATGATGAATGTGGAATTTTCAGAAAAGAGATGTCTGCTTTGG 1535
QY	599	TTCAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCCTCTGTTTTCTAAGATTATCC 658
DB	1536	TTTAACTCTCTCTGGAGAGGATTACAAATTTGAAACTGGAACCTCTTCATCCTTATAATAC 1595
QY	659	CTGAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCTTGCATAAGCTG 718
DB	1596	CAGAACAGAGCACGTTTAAAGTAGTCTTTCAACAAGAGATTGAAATTTAACTGAAAAGCCAG 1655
QY	719	AGCAGGTGACATGCACAAACCCCTGGATTATAGTGAAGAGACCAAGAGACTGTTCCCCAGAAGA 778
DB	1656	AGGCTGTGAGATGGGAAAAGCTAGA-----GGGGCAAGGAGATGTGCCTACGCCAAAAC 1709
QY	779	TAAAGCAGCGCAGCTGAAACAGCCCGCAAGACCTTCATACCCTATTCAAGGCGAAGAAAG 838
DB	1710	AATTCGTAGCAGATGTAAGAAACCTATATCCATCATCATCTCCTT-----ATACAAGAA 1763
QY	839	ACTGGGTAACCTGGAGCTGAACTCAAAAGGAGGAGAGGAAAGAAAACCTTGATGGTG 898
DB	1764	ATTGGGATAAATTGGTTGGTGAGATCAAGAGAGAGAAAGATGAAGATTTGGAGGGAG 1823
QY	899	ATGCTGCATTGAACAAATCTTCCTCGTGACATCTACAAGAGATGCTGATGAAGATATCGCGA 958
DB	1824	ATGCAGCTTTAAACAGATTAATTTTCAGCAGATCTATTTCAGATGGTTCTTGATGAAGTGAAC 1883
QY	959	GGGCCATGATGAAGTCAATTCGTGGAATCAATGCGACTGTTCTCTCAACCAATTTGGAAG 1018
DB	1884	GTGCCATGAACAAATCTTTATGGAGTCGGGTGGTACAGTTTTTCAGTACCAACTGGTCTG 1943
QY	1019	ATGTTGGAGCAAGAAGGTAGAAGGGAGCCCCCTGATGGTATGGAGCTCAAGAAGTGGG 1078
DB	1944	ATGTAGTAAAGAGAAAGTTGAATCAATCTCTCTGATGATATGAATGGAAAAAGTACT 2003
QY	1079	AA 1080
DB	2004	AA 2005

RESULT 3
 US-09-248-796A-2330
 ; Sequence 2330, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDID
 ; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725

```
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2330
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-2330

Query Match          7.0%; Score 76.2; DB 4; Length 198;
Best Local Similarity 64.4%; Pred. No. 6.9e-11;
Matches 114; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 910 AACAAATCTTCGTGACATCTACAGGATGCTGATGAAGATATCGGAGGCGCATATG 969
Db 22 AATGACTTTTCCGGAATAATTCAAAGATGTTGACGAAGATCTAGAGAGCGCATATG 81

QY 970 AAGTCATTCTGGAATCAAAATGGCACTGTTCTCAACCAATTCGAAAGATGTTGGAGCA 1029
Db 82 AAGATTATGTTCAATCGAATGGAACTGTTTGACTCAAGTGGGATGAAGCTAAAGAT 141

QY 1030 AAGAAGCTAGAAGGGAGCCGCCCTTGATGGATGCTCAAGAAGTGGGAATACAA 1086
Db 142 AAAGAATTTGAGGTTTTTACCACCACAGATGGATGGAAGTCAAGAAATGGATACGTA 198

RESULT 4
US-09-949-016-1499
; Sequence 1499, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1499
; LENGTH: 1735
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1499

Query Match          6.8%; Score 74.2; DB 4; Length 1735;
Best Local Similarity 55.1%; Pred. No. 7.4e-10;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 GCGTCGATCTGGAGAGCAAGGCCAAGGAGCGCTTCTGTCGACGACGACTTCGAGCTGGCC 66
Db 82 GCAGAGGAGCTCAAGACTCAGGCCAATGACTTCTTAAAGCCAAAGGACTACGAGAACGCC 141

QY 67 ACCGAGCTCTACAGCCAGGCGCATCGAGCGGGCGCCGCCACCGCGGACTCTATGCGGAC 126
Db 142 ATCAAGTTCTTACAGCCAGGCGCATCGAGCTGAACCCAGCAATGCCATCTACTATGGCAAC 201

QY 127 CGGCCCGAGCGGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTTAAACAA 186
Db 202 CGAGCTCTGCTTACCTGCGCACTGAGTGTATGGCTACGGCTTGGAGAGCGCCACCGGG 261

QY 187 GCAATTGAGCTTGATCTCTATGATGCATAAAGCTTTACTACCGGAAAGGTGCTGCATGCA 246
Db 262 GCCATTGAGCTTGACAGCAAGAGTACATCAAGGGTTTATTACCGCGGGTGTCCAGCAATG 321

QY 247 AAGCTTGAAGATACCAAACTGC 269
```

```
Db 322 GCAGTGGCAAGTTCGCGGCGC 344

RESULT 5
US-09-016-434-1067
; Sequence 1067, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1067:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1982 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1122930
US-09-016-434-1067

Query Match          6.8%; Score 74.2; DB 4; Length 1982;
Best Local Similarity 55.1%; Pred. No. 7.9e-10;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 GCGTCGATCTGGAGAGCAAGGCCAAGGAGCGCTTCTGTCGACGACGACTTCGAGCTGGCC 66
Db 60 GCAGAGGAGCTCAAGACTCAGGCCAATGACTTCTTAAAGCCAAAGGACTACGAGAACGCC 119

QY 67 ACCGAGCTCTACAGCCAGGCGCATCGAGCGGGCGCCGCCACCGCGGACTCTATGCGGAC 126
Db 120 ATCAAGTTCTTACAGCCAGGCGCATCGAGCTGAACCCCAAGCAATGCCATCTACTATGGCAAC 179

QY 127 CGGCCCGAGCGGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTTAAACAA 186
Db 180 CGAGCTCTGCTTACCTGCGCACTGAGTGTATGGCTACGGCTTGGAGAGCGCCACCGGG 239

QY 187 GCAATTGAGCTTGATCTCTATGATGCATAAAGCTTTACTACCGGAAAGGTGCTGCATGCA 246
Db 240 GCCATTGAGCTTGACAGCAAGAGTACATCAAGGGTTTATTACCGCGGGTGTCCAGCAATG 299

QY 247 AAGCTTGAAGATACCAAACTGC 269
```

Db 300 GCACTGGGCAAGTTCGGGGCGC 322

RESULT 6
US-09-825-497A-40
; Sequence 39, Application US/09825497A
; Patent No. 659742
; GENERAL INFORMATION:
; APPLICANT: Honkanen, Richard E.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF HUMAN SERINE/THREONINE PRO
; FILE REFERENCE: ISPH-0572
; CURRENT APPLICATION NUMBER: US/09/825,497A
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 40
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-825-497A-40

Query Match 6.8%; Score 74.2; DB 4; Length 1982;
Best Local Similarity 55.1%; Pred. No. 7.9e-10;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 GCCTGGATCTGGAGAGCAAGCCAGGAGGCTTCTGTCGACGACGACTTCGAGCTGGCC 66
Db 60 GCAGAGGAGCTCAAGACTCAGGCAATGACTACTTCAAGGCCAAGGACTACGAGAACGCC 119
QY 67 ACCGAGCTCTACAGCCAGGCAATCGAGCGGGCCGCCACCGACCTCTATGCCGAC 126
Db 120 ATCAAGTTCTACAGCCAGGCAATCGAGCTGACCCAGCAATGCCATCTACTATGGCAAC 179
QY 127 CGCGCCAGCGGCAATCAAGCTCGGCAATACACTGAGGCTGTGGCGGATGCTAACAA 186
Db 180 CGCAGCTGCGCTACCTCGGCAATGAGTGTGCTATGGCTACGCGTGGGAGACGCCACGCG 239
QY 187 GCATTGAGCTGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
Db 240 GCCATTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
QY 247 AAGCTTGAAGAATACCAAACTGC 269
Db 300 GCACTGGGCAAGTTCGGGGCGC 322

RESULT 7
US-09-825-497A-39
; Sequence 39, Application US/09825497A
; Patent No. 659742
; GENERAL INFORMATION:
; APPLICANT: Honkanen, Richard E.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF HUMAN SERINE/THREONINE PRO
; FILE REFERENCE: ISPH-0572
; CURRENT APPLICATION NUMBER: US/09/825,497A
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 39
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-825-497A-39

Query Match 6.8%; Score 74.2; DB 4; Length 1987;
Best Local Similarity 55.1%; Pred. No. 7.9e-10;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 GCCTGGATCTGGAGAGCAAGCCAGGAGGCTTCTGTCGACGACGACTTCGAGCTGGCC 66
Db 65 GCAGAGGAGCTCAAGACTCAGGCAATGACTACTTCAAGGCCAAGGACTACGAGAACGCC 124
QY 67 ACCGAGCTCTACAGCCAGGCAATCGAGCGGGCCGCCACCGACCTCTATGCCGAC 126
Db 125 ATCAAGTTCTACAGCCAGGCAATCGAGCTGACCCAGCAATGCCATCTACTATGGCAAC 184
QY 127 CGCGCCAGCGGCAATCAAGCTCGGCAATACACTGAGGCTGTGGCGGATGCTAACAA 186
Db 185 CGCAGCTGCGCTACCTCGGCAATGAGTGTGCTATGGCTACGCGTGGGAGACGCCACGCG 244
QY 187 GCAATTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 246
Db 245 GCCATTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
QY 247 AAGCTTGAAGAATACCAAACTGC 269
Db 305 GCACTGGGCAAGTTCGGGGCGC 327

RESULT 8
US-09-517-779-1
; Sequence 1, Application US/09517779
; Patent No. 6660511
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Xu, Xiang
; TITLE OF INVENTION: Cell Cycle Proteins Associated with Rad9, Compositions
; FILE REFERENCE: A68293/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/517,779
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-517-779-1

Query Match 6.8%; Score 74.2; DB 4; Length 1987;
Best Local Similarity 55.1%; Pred. No. 7.9e-10;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 GCCTGGATCTGGAGAGCAAGCCAGGAGGCTTCTGTCGACGACGACTTCGAGCTGGCC 66
Db 65 GCAGAGGAGCTCAAGACTCAGGCAATGACTACTTCAAGGCCAAGGACTACGAGAACGCC 124
QY 67 ACCGAGCTCTACAGCCAGGCAATCGAGCGGGCCGCCACCGACCTCTATGCCGAC 126
Db 125 ATCAAGTTCTACAGCCAGGCAATCGAGCTGACCCAGCAATGCCATCTACTATGGCAAC 184
QY 127 CGCGCCAGCGGCAATCAAGCTCGGCAATACACTGAGGCTGTGGCGGATGCTAACAA 186
Db 185 CGCAGCTGCGCTACCTCGGCAATGAGTGTGCTATGGCTACGCGTGGGAGACGCCACGCG 244
QY 187 GCAATTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 246
Db 245 GCCATTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
QY 247 AAGCTTGAAGAATACCAAACTGC 269
Db 305 GCACTGGGCAAGTTCGGGGCGC 327

RESULT 9
US-09-513-999C-4027
; Sequence 4027, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.

; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-551-974A-101

Query Match 5.6%; Score 61.2; DB 4; Length 2735;
Best Local Similarity 52.3%; Pred. No. 3.2e-06;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 6 CGCGTCGGATCTGAGAGCAAGGCCAAGGAGGCGCTTCGTGCGAGCAGCACTTCGAGCTGGC 65
DB 1698 CGCGAGCAGAGAGAGACGAGGTAACCACTTCAAGGAGGATTAAGTTCCCGAGGC 1757

QY 66 CACCGAGCTCTACAGCGAGCCATCGACGCGCGGCGCCGCCACCGCGCACTCTATGCGGA 125
DB 1758 CGTGGCAGCGTACAGGAGGCCATCAAGCGCAACCCCTGCGGAGCACAACCTCTACAGCA 1817

QY 126 CCGCGCCAGCGCCACATCAAGCTCGCACTACACTAGGCTGTGGCGGATGCTAACAA 185
DB 1818 TCAGCGGCGCGGTACATCAAGCTTGAGGCTTCAACGAGCGCCCTCAAGGACGCGGAGAA 1877

QY 186 AGCAATTGAGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 245
DB 1878 GTGCATTGAGCTGAGCGGAGCTTTTGTAAAGGCTACGCGCGCAAGGGTCACTGCTTACTT 1937

QY 246 TAAGCTTGAAGATACCA 263
DB 1938 TTGACCAAGCAGTACAA 1955

RESULT 13
US-09-565-501A-101
; Sequence 101, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillion, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565.501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-565-501A-101

Query Match 5.6%; Score 61.2; DB 4; Length 2735;
Best Local Similarity 52.3%; Pred. No. 3.2e-06;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 6 CGCGTCGGATCTGAGAGCAAGGCCAAGGAGGCGCTTCGTGCGAGCAGCACTTCGAGCTGGC 65
DB 1698 CGCGAGCAGAGAGAGACGAGGTAACCACTTCAAGGAGGATTAAGTTCCCGAGGC 1757

QY 66 CACCGAGCTCTACAGCGAGCCATCGACGCGCGGCGCCGCCACCGCGCACTCTATGCGGA 125
DB 1758 CGTGGCAGCGTACAGGAGGCCATCAAGCGCAACCCCTGCGGAGCACAACCTCTACAGCA 1817

QY 126 CCGCGCCAGCGCCACATCAAGCTCGCACTACACTAGGCTGTGGCGGATGCTAACAA 185
DB 1818 TCAGCGGCGCGGTACATCAAGCTTGAGGCTTCAACGAGCGCCCTCAAGGACGCGGAGAA 1877

QY 186 AGCAATTGAGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 245
DB 1878 GTGCATTGAGCTGAGCGGAGCTTTTGTAAAGGCTACGCGCGCAAGGGTCACTGCTTACTT 1937

QY 246 TAAGCTTGAAGATACCA 263
DB 1938 TTGACCAAGCAGTACAA 1955

RESULT 14
US-09-639-206A-101
; Sequence 101, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillion, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639.206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-639-206A-101

Query Match 5.6%; Score 61.2; DB 4; Length 2735;
Best Local Similarity 52.3%; Pred. No. 3.2e-06;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 6 CGCGTCGGATCTGAGAGCAAGGCCAAGGAGGCGCTTCGTGCGAGCAGCACTTCGAGCTGGC 65
DB 1698 CGCGAGCAGAGAGAGACGAGGTAACCACTTCAAGGAGGATTAAGTTCCCGAGGC 1757

QY 66 CACCGAGCTCTACAGCGAGCCATCGACGCGCGGCGCCGCCACCGCGCACTCTATGCGGA 125
DB 1758 CGTGGCAGCGTACAGGAGGCCATCAAGCGCAACCCCTGCGGAGCACAACCTCTACAGCA 1817

QY 126 CCGCGCCAGCGCCACATCAAGCTCGCACTACACTAGGCTGTGGCGGATGCTAACAA 185
DB 1818 TCAGCGGCGCGGTACATCAAGCTTGAGGCTTCAACGAGCGCCCTCAAGGACGCGGAGAA 1877

QY 186 AGCAATTGAGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 245
DB 1878 GTGCATTGAGCTGAGCGGAGCTTTTGTAAAGGCTACGCGCGCAAGGGTCACTGCTTACTT 1937

QY 246 TAAGCTTGAAGATACCA 263
DB 1938 TTGACCAAGCAGTACAA 1955

RESULT 15
US-09-874-923-101
; Sequence 101, Application US/09874923
; Patent No. 6638517

```
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-874-923-101

Query Match          5.6%; Score 61.2; DB 4; Length 2735;
Best Local Similarity 52.3%; Pred. No. 3.2e-06;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Qy      6  CGCGTCGGATCTGAGAGCAGGCCAAGAGGGCCCTTCGCGACGACGACTTCGAGCTGGC 65
Db      1698  CGCGAAGCAGAAAGAACGAGGTAAACCGTACTTCAAGGAGGATAAGTTCCCGAGGC 1757

Qy      66  CACCGAGCTCTACAGCCAGGCGCATCGACGCGGGCCGCCACCGCCGACCTCTATGCCGA 125
Db      1758  CGTGGCAGCGTACACGGAGGCCATCAGCGCAACCTGCCGAGCACACCTCCTACAGCAA 1817

Qy      126  CGCGCGCCAGCGCACATCAAGCTCGGCAACTACATGAGGCTGTGGCGGATGCTAACAA 185
Db      1818  TCGCGCGCCGCGTACATCAAGCTTGGAGCCTTCAAGCGCCCTCAAGGACGCGGAGAA 1877

Qy      186  AGCAATTGAGCTTGATCCTATGATGCATTAAGCTTACTACCGGAAAGGTGCTGCATGCAT 245
Db      1878  GTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACCGCGCAAGGGTCATGCTTACTT 1937

Qy      246  TAAGCTTGAAGAATACCA 263
Db      1938  TTGACCAAGCAGTACAA 1955
```

Search completed: April 23, 2005, 08:59:04
Job time : 195.463 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:35:27 ; Search time 4819.69 Seconds
(without alignments)
10918.214 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggccgcgtcgatctgga.....tcaagaagtgggaataactaa 1086

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ets.*
12: gb_ey.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734.2	67.6	1409	8 AF439974	AF439974 Hordeum v
2	733.6	67.6	1222	8 AF192467	AF192467 Oryza sat
3	456.2	42.0	984	8 RID251317	AJ251317 Rubus ida
4	454.8	41.9	1113	8 AF494083	AF494083 Nicotiana
5	419.2	38.6	1366	8 AY085464	AY085464 Arabidops
6	418.2	38.5	1354	8 ATH318019	AJ318019 Arabidops
7	417.6	38.5	1077	6 AX412470	AX412470 Sequence
8	417.6	38.5	1108	8 AY150487	AY150487 Arabidops
9	417.6	38.5	1290	8 AF439976	AF439976 Arabidops
10	417.6	38.5	1400	8 AF428340	AF428340 Arabidops
11	417.6	38.5	1420	8 AF370229	AF370229 Arabidops
12	413.8	38.1	1197	8 AJ620883	AJ620883 Brassica
13	407.2	37.5	1068	8 AJ620882	AJ620882 Brassica
14	394.8	36.4	1291	8 AF439975	AF439975 Arabidops
15	393.2	36.2	1084	8 AY096395	AY096395 Arabidops
16	393.2	36.2	1348	8 AY063978	AY063978 Arabidops
17	285.8	26.3	968	8 ATH318020	AJ318020 Arabidops
18	260.4	24.0	582	8 AK059268	AK059268 Oryza sat
19	207.8	19.1	634	8 AF516180	AF516180 Nicotiana

20	207.8	19.1	634	8	AF516181	Nicotiana
21	164.4	15.1	272	6	AR249214	Sequence
22	159.8	14.7	1606	5	AJ720706	Gallus ga
23	156.4	14.4	1002	6	CQ724202	Sequence
24	156.4	14.4	1002	9	AF132856	AF132856 Homo sapi
25	156.4	14.4	1025	6	BD209970	Human tra
26	156.4	14.4	1235	9	AY321358	Homo sapi
27	156.4	14.4	1535	9	BC000911	Homo sapi
28	156.4	14.4	1555	9	HSA344097	Homo sapi
29	156.4	14.4	1776	6	CQ490463	Sequence
30	156.4	14.4	1776	6	CQ496307	Sequence
31	156.4	14.4	2053	6	BD192630	123 human
32	156.4	14.4	2053	6	AR184101	Sequence
33	156.2	14.4	1002	9	BT009798	Homo sapi
34	152.6	14.1	1518	5	BC072118	Xenopus l
35	146.6	13.5	139201	8	AP002970	Oryza sat
36	146.6	13.5	150022	8	AP003213	Oryza sat
37	145.2	13.4	1314	5	CR761938	Xenopus t
38	139.4	12.8	1104	10	BC009167	Mus muscu
39	137.6	12.7	153865	8	AP002744	Oryza sat
40	119	11.0	1089	9	AF068289	Homo sapi
41	114.6	10.6	154558	8	AC107206	Oryza sat
42	111	10.2	137843	9	AC138972	Homo sapi
43	107.8	9.9	304	4	BT344098	Bos tauru
44	90.8	8.4	1242	6	AX488766	Sequence
45	89	8.2	110000	8	CR382128_09	Continuation (10 o

ALIGNMENTS

RESULT 1
AF439974
LOCUS AF439974 1409 bp mRNA linear PLN 06-NOV-2003
DEFINITION Hordeum vulgare SGT1 (SGT1) mRNA, complete cds.
ACCESSION AF439974
VERSION AF439974.1 GI:17017305
KEYWORDS
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 1409)
Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A.,
Shirasu, K. and Schulze-Lefert, P.
The RAR1 interactor SGT1, an essential component of R
gene-triggered disease resistance
Science 295 (5562), 2073-2076 (2002)
JOURNAL MEDLINE
PUBMED 21893744
REFERENCE 2 (bases 1 to 1409)
Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A.,
Shirasu, K. and Schulze-Lefert, P.
Direct Submission
Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre,
Colney Lane, Norwich NR47UH, UK
FEATURES
source
1. .1409
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
1. .1409
/genes="SGT1"
9. .1130
/genes="SGT1"
/codon_start=1
/product="SGT1"
/protein_id="AAL33610.1"
/db_xref="GI:17017306"
translations="MAAASDLESKAKEAFVDDDFELAELVTOATEAGPATAEALYA
DRAQHKILGSYTEAVADANKAIELDPSMHKAVLRKGSACIKLEPYOTAALRVGSS
YASGDRFRFLMECDRIAEASQAPVKVAAAANAAPATSSGATTVTTEADQDGENM


```

126 GCGCGGAGCTCTACACGAGCAATCGAGCCGACCCCGCCACCGCGAGCTCTACGCC 185
124 GACCGCGCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTTAAC 183
186 GACCGCGCCAGGCGCCATATCAAGCTAGCAACTACACTGAGGCTGTAGCTGATGCTAAC 245
184 AAAGCAATTGAGCTTGATCTCTATGATGATCAATAAGCTTACTACCGGAAAGGTGCTGCATGC 243
246 AAGGCCATTGAACCTTGACCCATCAATGACAAAGGCTTATCTTCGTAAGGCGCTGCATGT 305
244 ATTAAGCTTGGAAGATACAAACTGCAAGGCTGCTCTTGAGTTGGGTCTTCTTATGCA 303
306 ATACGACTGGAGGAGTATCAACTGCAAAAGCAGCTCTTGAATGGGTACTCGTTCGCA 365
304 TCAGCGGATTCAGGTTTGCTGCTTATTTGAAGGATGTGATGAGCGCATCGCTGAGGAA 363
366 TCTGGTACTCAAGGTTTACTCGCTAATGAAGGAGTGTGATGAGCGCATTCGCTGAGGAG 425
364 TCTAGCCAGGCGACAGTAAAGAAATGTTGAGG-----CTACTGTGGCTGCT 408
426 CTTACTAGTCCCTTTAAGAGCTGAAGATGGAGCGCTGCGCCCTCTGTGCTTCT 485
409 ACTATTGAGGCAAGGAGGATTTCAAAATATGGAGATACACCAAGCTGATAGAACCC 468
486 TTTGTTGAGGAAAGGATGCTGCAAAACATGGATATACACCAAGTGTAGT----- 541
469 CCAAGCAAAACCAAAATACAGGATGACTATACAAAGTGCACAGAGTGGTACTGACA 528
542 --AGTGAAGCCAAATACAGGACGACTTCTACAAACAGTGTGATGAGAGTGTATGACA 599
529 ATATTGCTAGGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 588
600 ATTTTGAAGGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 659
589 AGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 648
660 AGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719
649 AAGATATTCCTGAGAAATGCAATATCAAGTCTTATCCCAAGGTTGAAATACGCTT 708
720 AAGATATTCCTGAGAAAGCAGATACCAAGTGTATCCAGCAAGGTTGAAATACGCTG 779
709 GCAAAAGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 768
780 GCTAAAGCTGACAGATTTACATGAGACCTTCTGATGATGATGATGATGATGATGATG 839
769 CCCAGAGATTAAGCAGCGAGCTGAAACAGCCCGCCAGACCTTCAATCCCATCTTCAAAG 828
840 CCACAAAGATATCCCTCCAGCTGATATCGGCCAGAGGCCATCATATCTTCTCTCAAA 899
829 GCGAAAAAAGCTGGGATTAATCTGGAAGCTGAAGTCAAAAAAGGAGGAGGAGGAGAA 888
900 TCCAAGAAAGACTGGGATTAATCTGGAAGCTGAAGTCAAAAAAGGAGGAGGAGGAG 959
889 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948
960 CTTGAAGCGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019
949 GATATGCGGAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1008
1020 GACATGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1079
1009 AATTGGAAGATGTTGGAGCAAGAGGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1068
1080 AATTGGAAGATGTTGGCTGCAAGAGGATGAAGGAGGAGGAGGAGGAGGAGGAGG 1139
1069 AAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
1140 AAGAAATGGGAGTACTTA 1157

```

RESULT 3

RID251317
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

RID251317
Rubus idaeus partial mRNA for putative protein phosphatase.
AJ251317
AJ251317.1 GI:6468694
protein phosphatase.
Rubus idaeus
Rubus idaeus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rubus.

Schroeder, G., Lurz, G. and Schroeder, J.
cDNA for protein phosphatase like protein from raspberry (Rubus
idaeus)
Unpublished
2 (bases 1 to 984)
Schroeder, G.
Direct Submission
Submitted (06-NOV-1999) Schroeder G., Institut fuer Biologie II,
Universitaet Freiburg, Schaezlestr. 1, Freiburg D-79104, Germany

Location/Qualifiers
1..984
/organism="Rubus idaeus"
/mol_type="mRNA"
/db_xref="taxon:32247"
<1..984
/codon_start=1
/product="putative protein phosphatase"
/protein_id="CAB61630.1"
/db_xref="GI:6468695"
/db_xref="UniProt/TREMBL:Q9SM53"
/translation="PKNSLYSDRAQNMKSNLTAEVADANKAIELDPSLFKAYLRK
-GIACMLSEYRTAKAALGASLAPGSRFAKMKSCDELTAENGPPKQPMETTT
EIVAEDEVPPSNVETVAPVKYRHEFYQAEVVTIFAKIGIAPENVAIVDFGPO
ILSVSDVPGSVHIFOPRLFGLIKPEKRFELVSTVEIRLAKAEPNHWTSLEBPKD
NPVPLKYNAPVIGAQRPSPSKPRVDMDKLEQVKKEKDEKLDGDAALNKPFQDI
YKDAEDTRRAMRKSFVESNGTVLSTNWKVGNKKVSGSPDGMEMKKWEI"

source
1..984
/organism="Rubus idaeus"
/mol_type="mRNA"
/db_xref="taxon:32247"
<1..984
/codon_start=1
/product="putative protein phosphatase"
/protein_id="CAB61630.1"
/db_xref="GI:6468695"
/db_xref="UniProt/TREMBL:Q9SM53"

CDS

ORIGIN

Query Match 42.0%; Score 456.2; DB 8; Length 984;
Best Local Similarity 68.5%; Pred. No. 7.9e-72;
Matches 677; Conservative 0; Mismatches 303; Indels 9; Gaps 3;
Qy 100 CCGCCACCGCCGACCTCTATGCGGACCGGCCCGGCGCACATCAAGCTCGGCAACTAC 159
Db 1 CCCAAGAACTCCGAGCTCTACTCCGACCGGCTCAGGCCAACATGAAATCAACAATCTC 60
Qy 160 ACTGAGGCTGCGCGGATGCTTACAAAGCAATTCAGCTTGATCTATGATGATGATGATG 219
Db 61 ACCGAGGCTGTGCGGATGCGAACAGGCCATTGAGTTGGACCCGCTCTTTGTTCAAGCC 120
Qy 220 TACTACCGGAAAGGTGCTGATGATTAAGCTTGAAGAAATACCAAACTGCAAAAGGCTCT 279
Db 121 TACTTGGCAAGAAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 280 CTTGAGTTGGGTTCTTCTTATGATCAGGCGATTCAGAGTTTGTCTGCTTATGAGGAA 339
Db 181 TTGGAATCTGCTGCTCTTCTTGGCACCGAGGAAATCCAGATTCGCTTAAATGATAAAGG 240
Qy 340 TGTGATGAGCGCATCGCTGAGGATCTAGCAGGACCAAGTAAAGATGTTGAGGCTACT 399
Db 241 TGTGATGAGCTCATTTGAGAGGAAATTTGGTGGAGCCACC--GAGCAACCAATGGAACA 297
Qy 400 GTGGCTGCTACTATTGAGGACCAAGGAGGATTTCAAAATATGGAGAAATACACCAAGG 459
Db 298 ACAACTACAGAGATGTTGAGAGATGTTGAGCCAGTGGACCCGCTTCCATGAGGTG 357
Qy 460 ATAGAACCCCAAGCAAAATACAGGATGATGATGATGATGATGATGATGATGATGATG 519
Db 358 ACAGTAGCCTGTGCAAAACCAAAATACAGGATGATGATGATGATGATGATGATGATG 417
Qy 520 GTACTGCAATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579


```

418 Db GTTGTGACTATATTGCAAGGGCATACCTGCCGAAATGTTGCTGTGATTTTGGTCCA 477
580 QY CAGATGTTGATGATCCATTGAATGTTCTGGTGAAGAACCATACCATTTTCAGCCCGCT 639
478 Db CAAATATTAAAGTTCAGCATTTGATTTTCCCGGTGAAGTACATATACATTTTTCAGCCCTGC 537
640 QY CTGTTTCTTAAGATTATCCCTGAGAAATGCAAAATCAAGTCTTATCCACCAAGTTGAA 699
538 Db CTATTGGAAGATCATCTCTGAAAGTGAGATTGAGGTTTGTCTACCAAGTTGAA 597
700 QY ATACGCCCTTGCAAAAGCTGACAGGTGACATGGAACCAACCTGGATTATAGTGAAGACCA 759
598 Db ATTGCGCTTGCAAAAGCTGAACCTAATCACTGGACTTCTCTGAGTTCA---GCAAGGAT 654
760 QY AAGACTGTTCCCAAGATGAAGCAGCCGAGTGAACAGCCGCAAGACCTTTCATACCCA 819
655 Db AACCCCTTCTCTTAAAGGTTAATGCCCCGATTTATGGAGCCCAAGGCCATCTTACCCA 714
820 QY TCTTCAAAAGGCGAAA---AAGACTGGGATAAAGCTGGAAGCTGAAGTCAAAAAGGAGGAG 876
715 Db TCCTTCAAAACCAAAAGGTTAGATTGGGACAAAGCTTGAAGCCCAAGTGAAGAGGAGAA 774
877 QY AAGGAAGAAAACCTTGATGATGATGCTGCAATGAACAAAATCTTCCGTGACATCTACAAG 936
775 Db AAGATGAGAAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
937 QY GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
835 Db GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
997 QY GTTCTCTCAACCAATTCGAAGATGTTGGAGCAAGAGGTTAGAGGAGGAGGAGGAGGAGGAG 1056
895 Db GTGCTTCAACCAACTGGAAGAGTGGGACAAAGAGTGGGACAAAGAGTGGGAGGAGGAGGAG 954
1057 QY GGTATGAGCTCAAGAAAGTGGGAATCTA 1085
955 Db GGTATGAGATGAAGAAATGGGAGATCTA 983

RESULT 4
AF494083 LOCUS Nicotiana benthamiana SGT1 mRNA linear PLN 30-OCT-2003
DEFINITION Nicotiana benthamiana SGT1 mRNA, complete cds.
ACCESSION AF494083
VERSION AF494083.1 GI:29468338
KEYWORDS
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
REFERENCE 1. (bases 1 to 1113)
AUTHORS Liu, Y., Schiff, M., Serino, G., Deng, X.W. and Dinesh-Kumar, S.P.
TITLE Role of SCF ubiquitin-ligase and the COP9 signalosome in the N
gene-mediated resistance response to Tobacco mosaic virus
JOURNAL Plant Cell 14 (7), 1483-1496 (2002)
MEDLINE 2211573
PUBMED 12119369
REFERENCE 2. (bases 1 to 1113)
AUTHORS Liu, Y., Schiff, M., Serino, G., Dinesh-Kumar, S.P. and Deng, X.-W.
TITLE Direct Substitution
JOURNAL Submitted (22-MAR-2002) MCDB, Yale University, New Haven, CT 06520,
USA
FEATURES
source
1. .1113
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
1. .1113
/codon_start=1
/product="SGT1"
/protein_id="AA085509.1"
CDS

```

```

/db_xref="GI:29468339"
/translation="MASDLETRAKAEAFIDHFEFELAVDLYTOAIFAMTPKNAELFADRAQ
ANIKNYFEAVVDANKATELDSPMSKAYLRKGLAKMLEYQTAKAALLETGASLAPA
ESRFTKLIKECDRIAEAGELPNOSVDKTSNVVAPPVASELSDNVAAPKQADPTVN
LSYQSSAARPKYRHFEPYOKPEEVVTIFAKGI PAKNVVVD FGEIILSVISIDVPGDITY
SFQPRFUGKITPAKRIYENVSTKIEIRLAKAEPLHWTSLYTRASAVORPNVSSAP
RFSYSSKLVHDWDKLEAEVKKBEKDKLDGDAALNKFFRDYIKDADSDTRAMWKS
FVESNGTVLSTNWKVGTGTRKVEGSPDPGMELKKWEI"
ORIGIN
Query Match 41.9%; Score 454.8; DB 8; Length 1113;
Best Local Similarity 65.8%; Pred. No. 1.4e-71; Indels 36; Gaps 3;
Matches 729; Conservative
QY 7 GCCTCGGATCTGAGAGCAAGGCCAAGGAGGCTTCGTGACGACCACTTCGAGCTGGCC 66
Db 4 GCCTCGATCTGAGACTAGGCTTAAGAGCTTTCATTGACGACCACTTTGAGCTGCC 63
QY 67 ACCGAGCTCTACAGCAGGCCATCGACGCGGGCCGCCCAACCTCTATGCGGAC 126
Db 64 GTTGACCTTTACACTCAAGCAATTGCGCATGACTCTTAAGAACTGAGCTTTTCGCGC 123
QY 127 CGCGCCAGCGGCACATCAAGCTCGGCAACTACACTGAGCTGTGCGGATGCTACAAA 186
Db 124 CGTGCTCAGGCCCAACATCAAACTCAACTTCTTCACTGAAGCTGTTGTTGATGCGAAC 183
QY 187 GCAATTTGAGCTTGATCCTATGATGATCAAAAGCTTACTACCGGAAAGGTGCTGATGCA 246
Db 184 GCAATTTGAGTTAGTCTTCAATGTCAAAAGCATATTTGCGTAAGGGTTGGCTGTATG 243
QY 247 AAGCTTTGAAGAAATACAAAAGCTGCTCTTGAAGTGGGTTCTTCTTATGATCA 306
Db 244 AAGCTTTGAAGAGTACAAAAGCTGCTTGAAGTGGGTTCTTCTTATGATCA 303
QY 307 GCGGATTTCAAGGTTTCTCTCTTATTTGAAGGATGATGAGCGCATCGCTGAGGATCT 366
Db 304 GCAGAGTCAAGGTTTCAAAAGTTAATCAAAAGATGATGAGCGCATTTGAGAGGAGCT 363
QY 367 AGCAGGCGCACGTAAG-----AATGTTGAGGCTACTGTGGCT 405
Db 364 GGAAGTACTTAATCAGTCGGTGGATAAAACCTCGGAAATTCGTAGCTCCCTGCA 423
QY 406 GCTACTATTGAGCAAGGAGGATTTCAAAATATGAGAGATACACCAAGTGA----- 460
Db 424 TCTGAGTCTTTGGACAATGTTGCTGTTGCGCTTAAAGATGCTCAACCACTGTCAACCTG 483
QY 461 -----TAGAAGCCCAAGCAAAATACAGCGATGACTACTACACAGTGCCACA 513
Db 484 TCCTATCAAGGATCTGTCGACAGCAAAATACAGGATGAATTTTACAGAGCCAGAG 543
QY 514 GAAAGTGTACTGACAAATATTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAATGATTTT 573
Db 544 GAGTGTGTGACTATATTGCAAGGGAATACAGCCAGAAATGTTGTTGTTGACTTT 603
QY 574 GGTGAACAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
Db 604 GGTGAACAAATACCTTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 663
QY 634 CCCGCTGTGTTTCTAAGATTATCCCTGAGAAATGCAAAATATCAAGTCTTATCCACCAAG 693
Db 664 CTTAGGTTGTTTGGCAAGATTAACCTGCAAAATGAGATATGAAGTGTGTACCAAA 723
QY 694 GTTGAATATAGCTTGCAGAAAGCTGAGCAGGTGACATGGAACACCTCGGATATAGTGA 753
Db 724 ATTGAGATCGCTTGCAGAAAGCTGAACCTTACACTGGACATCTCTTGAATATA---CG 780
QY 754 AGACCAAGATGTTTCCCGAAGATGAAGCAGCGAGCTGAAACAGCCCCCAAGACCTTCA 813
Db 781 AGAGCGTCTGCTGTAGTGCAGAGGCGCTTAATGTGTCTCATGATGCTCCCGCCCGCTAT 840
QY 814 TACCCATCTTCAAGGCGGAAAAGACTGGGATTAATACTGGAAGTGAAGTCAAAAAGGAG 873
Db 841 CTTTCTCGAAATTTGACATGTTGGATTAATATTAGAGCTGAAGTGAAGGAG 900

```

```

Qy 874 GAGAGGAGGAGAACTTGATGGTGTGCTGATTCATTCGACAAATCTTCCTGGTGCATCTAC 933
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 GAAAAGACGAGAAATTTGGATGGAGATGCAGATTCGACAAATTTTCCGAGACATTTAC 960
Qy 934 AAGGATGCTGATGAAGATATCGGAGGCGCATGATGAAGTCATTCGTGGATCAAAATGGC 993
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
961 AAAGATCCGATGAGGACACAGAGGACCATGATGAATCTTTGTGGATCTAATGGG 1020
Qy 994 ACTGTTCTCAACCAATTTGAAAGATGTTGGAGCAAGAAAGATGAAGGAGGCCCCCT 1053
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1021 ACTGTTTGTGCAAAATCGAAAGATTTGGCACAAGAGGTGCAAGGAGGCCCTCCG 1080
Qy 1054 GATGCTATGAGCTCAAGAGTGGGAATCTA 1085
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1081 GATGCTATGAGCTCAAGAGTGGGAATCTA 1112

```

```

RESULT 5
LOCUS AV085464 1366 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 152712 mRNA, complete sequence.
ACCESSION AV085464
VERSION AV085464.1 GI:21404174
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

```

```

REFERENCE
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troughan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome
annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22088475
PUBMED 12093376

```

```

REFERENCE
AUTHORS Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE
AUTHORS Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA

```

```

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.

```

```

FEATURES
source
1..1366
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="152712"
57..1133
/codon_start=1

```

CDS

```

/product="SGT1a"
/protein_id="AA062690.1"
/db_xref="GI:21553597"
/translaton="MAKELAKEAKEAFLDDDFVAVDLYSKAIDLDPNCAAFPADRAQ
ANTKNFTAEVVDANKALEPTEKAYLRKGTACMKLEYSYAKALEKAGSAVEN
EPIFKMIDECDLRIAEKEDLVQMPSPSPSSSTPLATEADAPPVIPAAPKAPMF
RHEFYKPEBAVVTIFAKVPKENVTVEFGEQILSVVIDVADGEAYHLPRLFKGIIP
EKCRFVLSITKVEIRLAKABIIITWASLEYKGGOSLVKPNVSSALSORPVVPSKPAK
DWKLEAEVKKQEKDEKLDGDAAMKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
WKEVGTGKKVSTPPDGMELKKWEY"

```

ORIGIN

```

Query Match 38.6%; Score 419.2; DB 8; Length 1366;
Best Local Similarity 63.1%; Pred. No. 3.4e-65;
Matches 681; Conservative 0; Mismatches 393; Indels 6; Gaps 2;

Qy 7 GCGTCGGATCTGGAGAGCAAGGCCAAGAGGCCCTTCGTCGACGACGATTCGAGCTGGCC 66
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
60 GCGAAGGAATTTAGCAGAGAAAGCTTAAAGAGCTTTTCTAGATGATGATTCGATGTTGCT 119
Qy 67 ACCGAGCTCTACAGCCAGGCCATCGACCGCGGGCCGCCACCGCGGACCTCTATGCCGAC 126
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 GTTGACTTATCTCCAAAGCCATTGACTTGGATCCCAATTTGCCCGGCTTTCTTCGCCGAT 179
Qy 127 CGCGCCCGAGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAACAA 186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 CGTGCTCAGGCCAACATCAAAATCGATACTTCACTGAAGCTGTGTAGATGCCAACAA 239
Qy 187 GCAATTGAGCTTGATCTTCATGTCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 GCCATTGAGTTGGAGCCAAAGCTTTGGCAAAAGCCCTATCTCAGAAAGGCACTGCTTGATG 299
Qy 247 AAGCTTGAAGATACCAAACTGCAAGGCTGCTCTTGAGTTGGTCTTCTTATGATCA 306
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 AAGCTAGAGAATATAGTACTGCTTAAAGCAGCCCTGGAAGGAGGAGCTTCTGTTGCCCG 359
Qy 307 GCGGATTTCAAGGTTTGTGCTGCTATTGAAGGAATGTGAGCGCATCGCTGAGGAATCT 366
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 AATGAACAAAGTTTAAAGAGATGATAGATGAATGCGATCTTCTGTTGCAAGAAAGAG 419
Qy 367 AGCCAGGACACAGTAAAGAAATGTTGAGGCTACTGTGCTGCTACTATTGAGGACAAGGAG 426
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 A---AAGATTTGGTTTCAAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 476
Qy 427 GATTTCAAAATATGGAGAATACACCAAGTAGTATAGAACCCCAAGCAACCAAAATAC 486
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 GCAACGGAAGCTGATGCTCTCTCTGTTTCAATTCCTGCAGCACCTGCGCAACCCAGTTC 536
Qy 487 AGCATGACTACTACACAGTGCACAGAGTGGTACTGACAAATTTGCTTAAGGTTGTT 546
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
537 AGGCAGAGTTCTTACCAGAAACAGAGAAGCGGTGGTGACAAATTTTCCGCAAAAAGTA 596
Qy 547 CTTGCTGATAGTGTAGTCAATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
597 CTTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTGTGTCATTGATGT 656
Qy 607 CTTGTTGAAGAACCATACCATTTTACGCCCGTCTGTTTTCTTAAGATATTCCTTGAGAAA 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTTGTTCCGGGAAGATAAATACCAGAGAAG 716
Qy 667 TGCANAATATCAAGCTTATCCACCAAGTTGAATAGCCCTTGCAAAAGCTGACAGCGGTG 726
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
717 TGCAGATTTGAAGTATTGTGCAAAAGTTGAGATCCGCTTGTGGAAGCAGAGATAATC 776
Qy 727 ACATGGACAACCTCGGATTTATAGTGAAGACCAAGACTGTTCGCCAAGAGATAAGCACG 786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
777 ACCTGGGCTTCCCTTGAATATGTAAGGGCAAGTGTTTTGGCCAAACCCATGTCTCA 836
Qy 787 CCAGCTGAAGAACGCCCCAAGACCTTATACCCATCTTTCAAGAGCGCAAAAAGACTGGAT 846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
837 TCAGC---GCTGTGCGCAGAGACCAAGTGTACCCATCTTCTTAAGCCAGCAAAAGACTGGAC 893
Qy 847 AAATGGAACTGAAGTCAAAAAGGAGGAGAGAGGAGAAAACCTTGTGTTGTTGCTGCA 906

```

894 AAGTTGAGCTGAGTGAAGAACAGAGAGGATGAGAGCTTGTGAGATGAGCT 953
 907 TTGAACAAATCTTCCTGATCATCTAAGAGTCTGATGAAGATATGCGAGGCGCATG 966
 954 ATGAACAAATTTTCAGCGACATATCTCGAGTGCAGATGAAGACATGAGCGGCAATG 1013
 967 ATGAGTCATTCGTTGGAATCAATGACCTGTTCTCTCAACCAATGGAAGATGTTGGA 1026
 1014 AACAAATCATTTGAGAGTGAATGGGACGGTACTGTCGACAACTGGAAGAGTTGGG 1073
 1027 GCAAGAGGTAGAGGAGCGCCCTCTGATGCTGATGAGCTCAAGAGTGGGAATCTAA 1086
 1074 ACTAAGAGGTGAGAGCATCTCCACGATGGCATGAGCTCAAGAGTGGGATTTGA 1133

RESULT 6

ATH318019 1354 bp mRNA linear PLN 23-OCT-2003
 Arabidopsis thaliana mRNA for SGT1-like protein (sgt1a gene).

DEFINITION Arabidopsis thaliana mRNA for SGT1-like protein (sgt1a gene).
 ACCESSION AJ318019
 VERSION AJ318019.1 GI:30524963
 KEYWORDS SGT1-like protein; sgt1a gene.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 Pawlowski, J., Holzmann, M., Berney, C., Fahrni, J., Gooday, A. J.,

Cedhagen, T., Habura, A. and Bowser, S. S.

The evolution of early Foraminifera

Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11494-11498 (2003)

22882917

PUBMED 14504394

2 (bases 1 to 1354)

Shen, W. H.

Direct Submission

Submitted (11-MAY-2001) Shen W. H., Biologie Cellulaire, IBMP/CNRS,

12, rue du General Zimmer, 67084 Strasbourg, FRANCE

Location/Qualifiers

1. 1354

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

1. 1354

/gene="sgt1a"

40. 1116

/gene="sgt1a"

/function="proteolysis"

/codon_start=1

/product="SGT1-like protein"

/protein_id="CAC85266.1"

/db_xref="GI:30524964"

/db_xref="UniProt/TREMBL:Q9SUT5"

/translation="MAKLEAKAEFLDDPFDVAVDLYSKAIDLPNCNAFFADRAQ
 ANIKDTEAVDANKAIELEPLAKYLKGTACKMLEVSTAKALEKASVAPN
 RPEKMDDEADLRIAEKDLVQPPSLPSSSTPLAETADAPPVPIPAAPAKWF
 RHEFKPEEAVTIFAKKPKENVTFVEGQILSVIVDVAEYAHQLPRLFGKIIP
 EKCFEVLSTKAEIRLAKAEITWASLEYKGSVLKPKNVSSALSQRVPTPSKPK
 DWDLKLEAVKQEKDEKLDGDAANKFFSDIYSSADEMRMANKSFAESNGTVLSTN
 WKEVTKKVESTPPDGMELKKWEY"

ORIGIN

Query Match 38.5%; Score 418.2; DB 8; Length 1354;
 Best Local Similarity 63.0%; Pred. No. 5.le-65;
 Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

7 GCGTCGATCTGAGAGCAAGCGGAGGCTTCTGTCGAGCGAGCTTCGAGCTGGCC 66
 43 GCCAAGGAATTAGCAGAGAGCTTAAGAGCTTTCTTAGATGATGATCTTGATGTTGCT 102
 67 ACCGAGCTCTACAGCGAGGCCATCGACGCGCGGCCGCCACCGCGACCTCTATGCGGAC 126

103 GTTGACTTATATCTCAAGGCCATTGACTTGGATCCCAATTGCGCGCCCTTCTTCGCCGAT 162
 127 CCGGCCAGCGGCACATCAAGCTCGGCACTACACTGAGGCTGTGCGGATCTTAACAAA 186
 163 CBTGCTCAGGCGCAACATCAAAATCGATACCTTCACTGAAGCTGTGTAGATCGGACAAA 222
 187 GCAATTGAGCTTGATCTCTATGATGATAAAGCTTACTACCGGAAAAGGTGCTGCATGCA 246
 223 GCATTTGAGTTGAGGCCAAGTTGGCAAAAGCTATCTCAGAAAGGCACTCTTGTGTATG 282
 247 AAGCTTGAAGAATACCAAAGCTGCTCTTGAAGTGGTTCCTTATGATGATCA 306
 283 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGAGCTCTCTTTCACCG 342
 307 GCGGATTCAAGCTTGTCTGCTCTATTGAGGAATGTGATGAGCGCATCGCTCAGGAATCT 366
 343 ATGAACCAAAAGTTTAAAGAGATGATAGATGAATCGATCTTCGTATTGCAAGAGAGAG 402
 367 AGCCAGGCAACAGTAAAGAAATGTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGAG 426
 403 A---AAGATTGGTTTCAGCGCATGCCACGAGTTTGGCTTCAAGCTCTACAACACCACTA 459
 427 GATTTCAAAATATGGAGAAATACACCAAGTATAGACCCCAAGCAAGCAACCAATAC 486
 460 GCAACGGAAGCTGATGCTCTCTCTGTTCCAAATTCCTGCGAGCACCTGCCAAACCGATG 519
 487 AGCATGACTACTTACAAACAGTGCCACAGAAAGTGTACTGACAAATATTGCTAAGGGTGT 546
 520 AGGCAAGATTCTACCAAGAACACAGAAAGCGGTGTGACAAATTTCCGCAAAAAGTA 579
 547 CTGCTGATAGTGTAGTCAATTGATTGTTGTAACAGATGTTGAGTATTCATTTGAAGTT 606
 580 CCTAAGGAGAACGTAACCTGTCGAGTTTGTGAGCAGATTTCTGAGTGTCTCATTTGATG 639
 607 CTGTTGAAGAACCATACATTTTCAAGCCCGTCTGTTTTCTAAGATTATCCTCGAGAA 666
 640 GCTGAGAGAGGAGCTTATCATCTCCAGCGAGATGTTCTCGGAAGATAATACAGAGAAG 699
 667 TGCAAAATATCAAGTCTTATCCCAAGGTTGAAATACGCTTGCAAAAAGCTGAGCAGGTG 726
 700 TGCAGATTGGAATTTGTCGACCAAGTTGAGTCCGTCTTTCGGAAGACGAGAGATAATC 759
 727 ACATGAGCAACCTCGGATTTATAGTGAAGACCAAGACTGTTCCCGCAGAGATAAGCAG 786
 760 ACNTGGGCTCCCTTGAATATGTTAAAGGCAAAAGTGTTTTGCCCAACCCCAATGTCTCA 819
 787 CAGCTGAACACAGCCCCCAAGACCTTATACCCATCTTCAAGGCGGAAAAGACTGGAT 846
 820 TCAGC---GCTGTCGAGAGACCAAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGAC 876
 847 AAATCGAAGCTGAAGTCAAAAAGGAGAGAGAAAGAAAACCTTGATGTTGATGTCGA 906
 877 AAGTTGAAGCTGAAGTGAAGAAACAGGAGAGAGTGAAGCTTGTGAGATGCACT 936
 907 TTGAACAAATTTCTTCGTCGATCACTCAAGGATGCTGATGAAGATATGCGAGGCGCATG 966
 937 ATGAACAAATTTTTCAGCGACATATATCTCGAGTGCAGATGAAGACATGAGCGCGCAATG 996
 967 ATGAAGCTATTCGTTGGAATCAAAATGCACTGTTCTCTCAACCAATTTGGAAGATGTTGGA 1026
 997 ACAAATCATTTGAGAGTGAATGGACGGTACTGTGCAAACTTGAAGAAGATTGGG 1056
 1027 GCAAGAGAGGTAGAAGGAGGCCCTCTGATGATGAGCTCAAGAGTGGGAATACTAA 1086
 1057 ACTAAGAAAGTGGAGAGCACTCCACGATGGCATGGAGCTCAAGAAAGTGGGAGTATGA 1116

RESULT 7

AX412470

LOCUS

DEFINITION Sequence 234 from Patent WO0222675.

ACCESSION AX412470

linear

1077 bp

DNA

PAT 14-JUN-2002

```

VERSION      AX412470.1  GI:21444928
KEYWORDS     Arabidopsis thaliana (thale cress)
SOURCE       Arabidopsis thaliana
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE    1
AUTHORS      Glazebrook, J., Wang, X., Dangl, J.L., Eulgem, T. and Zhu, T.
TITLE        Plant genes, the expression of which are altered by pathogen
              infection
JOURNAL      Patent: WO 0222675-A 234 21-MAR-2002;
              Syngenta Participations AG (CH); UNIVERSITY OF NORTH CAROLINA AT
              CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Xun (US); Dangl,
              Jeffrey L. (US); Eulgem, Thomas (US)
FEATURES     location/Qualifiers
              source      1..1077
                        /organism="Arabidopsis thaliana"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:3702"
ORIGIN
Query Match      38.5%; Score 417.6; DB 6; Length 1077;
Best Local Similarity 63.0%; Pred. No. 6.5e-65;
Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

Qy 7 CGCTCGATCTGGAGACGAGCCCAAGAGGCGCTTCGTGACGACGAGCTTCGAGCTGGCC 66
Db 4 GCCAAGGATTTAGCAGAGAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63
Qy 67 ACCGAGCTCTACAGCCAGGCGCATCGCGCGGCGCCGCCCGCCGACCTTATGCCGAC 126
Db 64 GTTGACTTATCTCAAGCCCATGATGTTGATCCCAATTGCGCGGCTTCTCGCCGAT 123
Qy 127 CGCGCCAGCGGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTTAACAAA 186
Db 124 CGTGCTCAGGCCACATCAAAATCGATACCTCACTGAGCTGTGTAGATGCGAACA 183
Qy 187 GCAATTGAGCTTGATCTATGATGCATAAAGCTTACTACCGAAAGTGTGTCATGCAATT 246
Db 184 GCCATTTGAGTTGGAGCCAAAGCTTGGCAAAAGCCCTATCTCAGAAAGGCGCACTGCTTGTATG 243
Qy 247 AAGCTTCAAGAAATACCAAACTGCAAAAGCTGCTCTGAGTTGGTCTTCTTATGCAATCA 306
Db 244 AAGCTAGAGAATATAGTACTGCTAAAGAGCGCCCTGGAAAGGAGGCTTCTGTTCACCG 303
Qy 307 GCGGATTCAGAGTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db 304 AATGAACCAAGTTTAAGAGATGATAGATGATGCGATCTTCGTATTGCAAGAGAG 363
Qy 367 AGCCAGGCACAGTAAAGAAATGTTGAGGCTACTGTGCTGCTACTTATTTGAGGACAGGAG 426
Db 364 A--AAGATTGGTTCAAGCGGATGCCACCGAGTTGCTTCAAGCTCTACAAACCACTA 420
Qy 427 GATTTCAATATGGAGATACACCAACAGTATAGATAGATAGATAGATAGATAGATAGAT 486
Db 421 GCAACGGAAGTGATGCTCTCTGTTTCAATTCCTCGAGCAGCTGCGCAACCGATGTT 480
Qy 487 AGGCATGACTACTACAAAGTGCACAGAGTGTGCTGACAAATATTTGCTAAGGGTGT 546
Db 481 AGGCAGAGTTCTACCAAGAACAGAGAGAGCGGTGTGACAAATTTTCGCAAAAGTA 540
Qy 547 CTTGCTGATAGTGPAGTCAATTGATTTTGGTGAACAGATGTTGAGTGTATCATTTGAAGTT 606
Db 541 CCTAAGGAGAAGCTAACTGTCGAGTTTGGTGAGCAGATTTCTGAGTGTGTCATTGATGTT 600
Qy 607 CTTGGTGAAGAACATACCATTTTCAGCCCGCTGTTTCTTAAGATATCCCTGAGAAA 666
Db 601 GCTGGAGAGAGGCTTATCATCTCCAGCCGAGATTTGTTGCGGAGAGATAATACCAAGAG 660
Qy 667 TGCATAATCAAGTCTTATCCACCAAGGTTGAATAGCGCTTGCAGAAAGCTGAGCAGGTG 726
Db 661 TGCAGATTTGAGTATTGTGACCCAAAGTTGAGATCCGCTTCTTGGCAAGCAGAGATATC 720

```

727 ACATGGCAACCCCTGGATTATAGTGAAGACCAAGAGCTGTTCCCGAGAGATGAAGCAG 786
 721 ACCTGGGCGCTCCCTTGAATATGTTAAAGGGCAAGAGTGTTTTGCCCAACACCAATGTCTCA 780
 787 CCAGCTGAAACAGACCCCAAGACCTTTCATACCCATCTTCAAGAGGCAAAAAGAGACTGGGAT 846
 781 TCAGC---GCTGTGCGAGAGACCAAGTGTACCCATCTTCTTAAGCCAGCAAAAGAGACTGGAC 837
 847 AAATGGAAGTGAAGTCAAAAAGGAGGAGAGAGAAAGAAACTTGTATGTTGATGCTGCA 906
 838 AAGTTGGAAGTGAAGTGAAGAAACAGAGAGAGAGTGAAGCTTGTATGAGAGATGACAGCT 897
 907 TTGAACAAATCTTCCGTGACATCTACAAGAGTCTGATGAAGATATGCGAGAGGCGCATG 966
 898 ATGAACAAATTTTTCAGCGACATATCTCGAGTGCAGATGAAGACATGAGGCGGCGCAATG 957
 967 ATGAAGTCAATCTGTTGAATCAAAATGCGCACTGTTCTCTCAACCAATTTGAAAGATGTTGGA 1026
 958 AACAAATCATTTGAGAGTGCAGATGCGATGGACGGTACTGTGACAAACTGGAAGAAGTTGGG 1017
 1027 GCAAGAAAGTAGAAGGAGGCGCCCTGATGTTGATGAGCTCAAGAAAGTGGGAATATCTAA 1086
 1018 ACTAAGAAAGTGGAGAGCACTCCACAGATGGCATGGAGCTCAAGAAAGTGGGAGTATTGA 1077

RESULT 8
 AY150487
 LOCUS
 DEFINITION Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete cds.
 ACCESSION AY150487 GI:23297701
 VERSION AY150487.1
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1108)
 Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Open Reading Frame (ORF) Clones
 Unpublished
 2 (bases 1 to 1108)
 Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT

Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.


```

contributed equally to this work as PIs.
FEATURES
  Location/Qualifiers
    1..1400
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /db_xref="taxon:3702"
      /chromosome="4"
      /clone="RAF109-50-H05 (R14746)"
      /ecotype="Columbia"
    1..131
    132..1208
      /note="unknown protein"
      /codon_start=1
      /product="At4g11260/F8L21_50"
      /protein_id="AA116270.1"
      /db_xref="GI:16226818"
      /translation="MAKELAEKAEFLDDDFVAVDLYSKAIDLDPNCAAFADRAQ
      ANIKDFTSAVDANDIAELEPTLAKYLRKGTACMKLEESTAKAALKEGASVAPN
      EPFKQMDICDLRIAEERDLVQPMPPSPSSSTPLATEADAPPVLPFAAPAKPMF
      RHFFYQKPEAVVTIFAKYKENVTFEFGQILSVVIDVAGEEAVHLQPLRFKLIIP
      EKCRFEVLSTKVEIRLAKABIIITWASLEYKGOSVLPKPNVSALSOPVYPSPSKPAK
      DWDLKLEAEVKQEKDEKLDGDAAMNKFSDIYSSADEDMRRANKNKSPAESNGTVLSIN
      WKVEYTKKVESTPPDGNELKWEY"
    misc_difference 703
      /note="compared to genome"
      /replace="g"
      1209..1400
    3'UTR
    ORIGIN
      Query Match      38.5%; Score 417.6; DB 8; Length 1400;
      Best Local Similarity 63.0%; Pred. No. 6.5e-65;
      Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

7  GCGTCGATCTGAGAGCAAGCCAGGAGCGCTTCGTCGACGACGACTTCGAGCTGCGC 66
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 GCCAAGGAATTAGCAGAGAAGCTTTCTAGATGATGACTTCGATGTTGCT 194
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ACCGAGCTCTACAGCCAGGCGCATCGAGCGCGCGCGCCGACCGACCTCTATGCGCGAC 126
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 GTTGACTTATCTCCAAAGCAATTGACTTGATGCCAATTCGCGCGCTTCTTCGCGCAT 254
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
127 CGCGCCAGCGGCACATCAGCTCGGCACTACATCAGAGCTGTGGCGGATGCTAACAAA 186
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 CGTCTCAGGCGCAACATCAAACTGATACTCACTGAAGCTGTGTAGATGCGCAACAAA 314
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
187 GCAATTGAGCTTGATCTATGATGATCAATAAGCTTACTACCGGAAAGTGTGATGATT 246
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 GCCATTGATTGGAGCCAACTTGGCAAAAGCCCTATCTCAGAAAGGCACCTGCTTGATG 374
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
247 AAGCTTGAAGAATACCAAACTGCAAGAGCTGCTCTTGAGTTGGGTTCTTCTTATGCAATCA 306
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 AAGCTAGAGAATATAGTACTGCTTAAGCAGCCCTGGAAAGGGAGCTTCTGTTGCAACCG 434
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
307 GGCATTCAGGTTGCTGCTCTATTGAGGAATGATGAGCGCATCGCTGAGGAATCT 366
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 AATGAACCAAAAGTTAAGAAAGATGATGATGAATGCGATCTTCGTATTGCAAGAAAGAG 494
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 AGCCAGGCGCCAGTAAGAAATGTTGAGGCTACTGTGCTGCTACTATTGAGGCAAGGAG 426
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 A---AAGATTGGTTACGCCGATGCCACCGAGTTGGCTTCAAGCTCTACAAACCACTA 551
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 GATTTCACAAATATGGGAATATACACCAAGTATAGAACCCCAAGCAAAACCAAAATAC 486
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
552 GCAAGGAAGTGTGCTCCTCTGTTCCAAATTCCTGCAGCACCTGCACCAACCGATGTT 611
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 AGGATGACTACTACAAAGTGCACAGAAAGTGTGATGACAAATATTGCTAAGGGTGT 546
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
612 AGGCACGAGTTCTACCGAAACCAAGAAAGACGGTGTGATGACAAATTTTCGCCAAAAGTA 671
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
547 CCTGCTGATGCTAGTCAATGATTGTTGGTGAACAGATGTTGAGTGTATCATTTGAAGTT 606
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
672 CCTAAGGAGACGTAACCTGTCGAGTTTGGTGGAGCAGATCTTGATGTTGCTGATGTT 731

```

```

Qy 607 CTTGGTGAAGAACCATACCATTTTTCAGCCCGCTGTTTCTTAAGATTATCCCTGAGAAA 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 732 GCTGGAGAGGAGCTTATCATCTCCAGCCGAGATTGTTGGGGAAGATAATACCAGAGAAG 791
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 667 TGCATAATATCAAGTCTTATCCCAAGGTTGAATAACGCTTCACAAAAGCTGAGCAGGTG 726
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 792 TGCAGATTGGAAGTATTGTCGCAAAAGTTGAGATCGTTCGTAAGCAGCAGAGATAATC 851
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 727 ACATGAGCAACCCCTGGATTATAGTGAAGACCAAGAACTGTTCCCGAGAGATAAGCAGC 786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 852 ACCTGGGCTCCCTTGAATATGTAAGGCAAAAGTGTTCGCCAAACCCCAATGTCCTCA 911
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 787 CGAGCTGAAGACAGCCCAAGACCTTCATACCACTTCCTCAAGCGGCAAAAAGACTGGAT 846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 912 TCAGC---GCTGTGCGAGAGCCAGGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGAC 968
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 847 ARACTGGAAGCTCAAGCTCAAAAGAGGAGAGAGAAAGAACTTGATGTGATGCTGTGCA 906
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 969 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAGAGATGAGAGCTTGATGAGATGCAAGCT 1028
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 907 TTGAACAAATCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGCGCATG 966
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1029 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGCGCGCAATG 1088
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 967 ATGAAGTCAATCTGGAATCAAAATGCGACTGTTCTCTCAACCAATGGAAGATGTTGGA 1026
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1089 AACAAATCATTTGCGAGTGAATGGAGCGGTACTGTGCAACAACTGGAAGAAGTTGGG 1148
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1027 GCAAGAGAGTGAAGGAGCGCCCTGATGATGATGAGCTCAAGAGTGGGAATACTAA 1086
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1149 ACTAAGAAAGTGGAGAGCACTCCACAGATGGCATGAGCTCAAGAGTGGGAGTATTGA 1208
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
AF370229 1420 bp mRNA linear PLN 18-SEP-2002
LOCUS Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete
DEFINITION cds.
ACCESSION AF370229
VERSION AF370229.1 GI:13877932
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1420)
AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabisdopsis Full Length cDNA Clones
TITLE Unpublished
JOURNAL 2 (bases 1 to 1420)
REFERENCE Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabisdopsis Full Length cDNA Clones
TITLE Direct Submission
JOURNAL Submitted (18-APR-2001) Plant Gene Expression Center, 800 Buchanan
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

```


The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Ban, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES

Location/Qualifiers

1..1420

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="RAPL04-14-007 (R09567)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector (Lambda Zap) as a XhoI/SstI insert."

gene

1..1420

/gene="At4g11260"

5'UTR

1..136

/gene="At4g11260"

misc_difference

32..34

/gene="At4g11260"

/note="not present in genomic sequence"

137..1213

/gene="At4g11260"

/codon_start=1

/evidence="experimental"

/product="unknown protein"

/protein_id="AAK4044.1"

/db_xref="GI:13877933"

/translation="MAKELAKEAKEFLDDFVAVLYSKAIDLDPNCAFFADRAQ ANKIDFTFAVDVAIAELEPTLAKAYLRKGTACMKLEYSYAKALEKGSVAEN EPKFKMIDIEEDLRIAEEDLVQPMPPSSSTTPLAYEADAPPVPIPAAPAKPMF RHBYQKPEAVRTIIPAKVPKENVTVEFGQIILSVVIDVAGEAYHLPRLFGKIIP EKCFEVLSTKVEIRLAKABIIITWASLEYKGQSVLPKPNVSSALSRLPVPSKPAK DWDKLEAVKQEKDEKLDGDAAWKKFSDIYSSADEMRMRMKNKSFAESNGIVLSTN WKVEGTIKKVESTPPDGMELKKWEY"

1214..1420

/gene="At4g11260"

3'UTR

ORIGIN

Query Match 38.5%; Score 417.6; DB 8; Length 1420;
Best Local Similarity 63.0%; Pred. No. 6.5e-65;
Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

7 GCGTCGGATCGGAGAGCCAGCCAGGAGCGCTTCGTCGACGACGCTTCGAGCTGCGC 66
|||||
140 GCCAAGGAATAGCAGAGAAGCTTAAGAAGCTTTCTAGATGATGACTTCGATGTGCT 199
|||||
67 ACCGAGCTCTACAGCAGCGCCATCGACCGCGCGCGCCGACCGCGACCTCTATGCCGAC 126
|||||
200 GTTGACTTATCTCCAAAGCATTGACTTGGATCCCAATTGCGCGCGCTTCTCGCGAT 259
|||||
127 CGCGCCAGCGGCACATCAAGCTGGCACTACGAGGCTGTGGCGGATGCTTAACAAA 186
|||||
260 CGTGCTCAGGCCAACATCAAAATGCATACTCACTGAAGCTGTGTGATGCGCAAAA 319
|||||
187 GCAATTGAGCTTGATCTCTATGATCATATAAGCTTACTACCGGAAGGTCGTCGATGATT 246
|||||
320 GCCATTGATTGGAGCCCAAGCTTGGCAAAAGCCCTATCTCAGAAAGGCGACTGCTGTATG 379
|||||
247 AAGCTTGAAGATACCAAACTGCAAGGCTCTCTTGAGTTGGGTCTTCTTATGCAATCA 306
|||||
380 AAGCTAGAGATATAGTACTGCTTAAGCAGCCCTGGAAAGGAGCTTCTGTGTCACCG 439
|||||

Qy 307 GCGATTCAGGTTTGTCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
|||||
Db 440 AATGACCAAGTTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
|||||
Qy 367 AGCCAGGACACAGTAAGAATGTTGAGGCTACTGTGGCTGCTACTATTATGAGGACAAAGG 426
|||||
Db 500 A---AAGATTGGTTGAGCGGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 556
|||||
Qy 427 GATTTCATAATATGGAGATATACACACAGTGTGATAGAACCCCAAGCAAAACCAAAATAC 486
|||||
Db 557 GCACGGAAGCTGATGCTCTCTCTGTTTCCCAATTCCTGCAGCACCTGCGCAACCCGATGTT 616
|||||
Qy 487 AGCATGACTACTACACAGTGCACACAGTGTGATGATGATGATGATGATGATGATGATGAT 546
|||||
Db 617 AGCAGGAGTTCTACCAAGAACCAAGAACGCGTGTGATGATGATGATGATGATGATGATGAT 676
|||||
Qy 547 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
|||||
Db 677 CCTAAGGAGAACGTAACGTCGAGTTTGGTGAGCAGATCTGAGTGTGATGATGATGATGAT 736
|||||
Qy 607 CTTGCTGAGAACCATACCATCTTTTTCAGCCCGCTCTGTTTCTTAAGATGATGATGATGATGAT 666
|||||
Db 737 GCTGAGAGGAGGCTTATCATCTCCAGCCGAGATTTGTTGCGGAAGATAATACCAAGAGAAG 796
|||||
Qy 667 TGCAAAATATCAAGTCTTATCCACCAAGTTGAAATACGCTTGCACCAAGCTGAGCAGGTG 726
|||||
Db 797 TGCAGATTGAGATTGTCGACCAAGTTGAGATCCGCTTTCGAAAGCAGAGATAATC 856
|||||
Qy 727 ACATGGAACACCCCTGGATTATAGTGAAGACCAAGAGCTGTTCGCCAGAGATGAAGCAG 786
|||||
Db 857 ACCTGGGCTCCCTTGAATATGTAAGGGCAAGTGTTCGCCCAACCCCAATGTCCTCA 916
|||||
Qy 787 CCAGCTGAAACAGCCCCCAAGACCTTTCATCCATCTTCAAAGGGCAAAAGAGACTGGGAT 846
|||||
Db 917 TCAGC---GCTGTGCGAGAGACCAAGTGTACCATCTTCTAAGCCAGCAAAAGAGACTGGGAC 973
|||||
Qy 847 AACTCGAAGCTGAAGTCAAAAGGAGGAGGAAGCAAGAACTTGTGATGATGATGATGATGAT 906
|||||
Db 974 AAGTTGAAGCTGAAGTGAAGAAACAGGAGAGGATGAGAAAGCTTGTGAGGAGATGACAGT 1033
|||||
Qy 907 TTGAACAAATCTTCCCTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCCATG 966
|||||
Db 1034 ATGACAAATTTTTCAGGACATATATCTCGATGTCAGATGAGACATGAGCGGGCAATG 1093
|||||
Qy 967 ATGAAGTCAATTCGCGAATCAATGCGACTGTTCTTCAACCAATGGAAGAGATGTTGGA 1026
|||||
Db 1094 AACAAATCAATTGCGAGAGTCGATGGAGCGGTACTGTGCAAACTGGAAGAAAGTGGG 1153
|||||
Qy 1027 GCAAGGAGGTAGNAGGAGCGCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1086
|||||
Db 1154 ACTAAGAAAGTGGAGAGCACTCCACAGATGGCATGGAGCTCAAGAAAGTGGGAGTATTGA 1213
|||||

RESULT 12

Qy 307 GCGATTCAGGTTTGTCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
|||||
Db 440 AATGACCAAGTTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
|||||
Qy 367 AGCCAGGACACAGTAAGAATGTTGAGGCTACTGTGGCTGCTACTATTATGAGGACAAAGG 426
|||||
Db 500 A---AAGATTGGTTGAGCGGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 556
|||||
Qy 427 GATTTCATAATATGGAGATATACACACAGTGTGATAGAACCCCAAGCAAAACCAAAATAC 486
|||||
Db 557 GCACGGAAGCTGATGCTCTCTCTGTTTCCCAATTCCTGCAGCACCTGCGCAACCCGATGTT 616
|||||
Qy 487 AGCATGACTACTACACAGTGCACACAGTGTGATGATGATGATGATGATGATGATGATGAT 546
|||||
Db 617 AGCAGGAGTTCTACCAAGAACCAAGAACGCGTGTGATGATGATGATGATGATGATGATGAT 676
|||||
Qy 547 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
|||||
Db 677 CCTAAGGAGAACGTAACGTCGAGTTTGGTGAGCAGATCTGAGTGTGATGATGATGATGAT 736
|||||
Qy 607 CTTGCTGAGAACCATACCATCTTTTTCAGCCCGCTCTGTTTCTTAAGATGATGATGATGATGAT 666
|||||
Db 737 GCTGAGAGGAGGCTTATCATCTCCAGCCGAGATTTGTTGCGGAAGATAATACCAAGAGAAG 796
|||||
Qy 667 TGCAAAATATCAAGTCTTATCCACCAAGTTGAAATACGCTTGCACCAAGCTGAGCAGGTG 726
|||||
Db 797 TGCAGATTGAGATTGTCGACCAAGTTGAGATCCGCTTTCGAAAGCAGAGATAATC 856
|||||
Qy 727 ACATGGAACACCCCTGGATTATAGTGAAGACCAAGAGCTGTTCGCCAGAGATGAAGCAG 786
|||||
Db 857 ACCTGGGCTCCCTTGAATATGTAAGGGCAAGTGTTCGCCCAACCCCAATGTCCTCA 916
|||||
Qy 787 CCAGCTGAAACAGCCCCCAAGACCTTTCATCCATCTTCAAAGGGCAAAAGAGACTGGGAT 846
|||||
Db 917 TCAGC---GCTGTGCGAGAGACCAAGTGTACCATCTTCTAAGCCAGCAAAAGAGACTGGGAC 973
|||||
Qy 847 AACTCGAAGCTGAAGTCAAAAGGAGGAGGAAGCAAGAACTTGTGATGATGATGATGATGAT 906
|||||
Db 974 AAGTTGAAGCTGAAGTGAAGAAACAGGAGAGGATGAGAAAGCTTGTGAGGAGATGACAGT 1033
|||||
Qy 907 TTGAACAAATCTTCCCTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCCATG 966
|||||
Db 1034 ATGACAAATTTTTCAGGACATATATCTCGATGTCAGATGAGACATGAGCGGGCAATG 1093
|||||
Qy 967 ATGAAGTCAATTCGCGAATCAATGCGACTGTTCTTCAACCAATGGAAGAGATGTTGGA 1026
|||||
Db 1094 AACAAATCAATTGCGAGAGTCGATGGAGCGGTACTGTGCAAACTGGAAGAAAGTGGG 1153
|||||
Qy 1027 GCAAGGAGGTAGNAGGAGCGCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1086
|||||
Db 1154 ACTAAGAAAGTGGAGAGCACTCCACAGATGGCATGGAGCTCAAGAAAGTGGGAGTATTGA 1213
|||||

LOCUS AJ620883 1197 bp mRNA linear PLN 16-JAN-2004

DEFINITION Brassica oleracea mRNA for SGT1-like protein (sgt1b-2 gene).

ACCESSION AJ620883

VERSION SGT1-like protein; sgt1b-2 gene.

KEYWORDS SGT1-like protein; sgt1b-2 gene.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE 1 Yemm, A.A., Holub, E. and Tor, M.

AUTHORS Unpublished

JOURNAL 2 (bases 1 to 1197)

REFERENCE Yemm, A.A.

AUTHORS Direct Submission

TITLE Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research,

JOURNAL Horticultural Research International, Wellesbourne, Warwick, UNITED

FEATURES	KINGDOM		Location/Qualifiers	
	source	1..1197	/organism="Brassica oleracea"	
gene	CDS	1..1197	/mol_type="mRNA"	
			/variety="Senna"	
gene	CDS	1..1197	/db_xref="taxon:3712"	
			/tissue_type="leaf"	
gene	CDS	1..1197	/gene="sgtlb-2"	
			/gene="sgtlb-2"	
gene	CDS	1..1065	/functions="role in disease resistance"	
			/codon_start=1	
gene	CDS	1..1065	/evidence="experimental"	
			/product="SGT1-like protein"	
gene	CDS	1..1065	/protein_id="CAF06581.1"	
			/db_xref="GI:40974917"	
gene	CDS	1..1065	/translation="MASELAEKAEFLDDPVDVADLYSKAIDLPSCASFFADRAQ ANIKLNFTEAVADNANKLEPTLAKAYLRKGAACMKLEAYATAKAALEKASVAPN ESKPEKIDECNLLIAEEKDLVQOVPTLPSSTTTPLAADAPPAPAKPMPRHEF YQPEEVGVALPAKGIKQNVNVEFGDQILSVVIDVAGEAVHFOPELFGK11PEKCR YEVLSUKLEAKAEIVTWASLEIGKQALLPKPNVASVRSORPVYPSKPKQKDWK LEAEVKQEKDEKLDGDAAMNKFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWEV GTKVVESTPPDGMELKKWEY"	
			/translating="MASELAEKAEFLDDPVDVADLYSKAIDLPSCASFFADRAQ ANIKLNFTEAVADNANKLEPTLAKAYLRKGAACMKLEAYATAKAALEKASVAPN ESKPEKIDECNLLIAEEKDLVQOVPTLPSSTTTPLAADAPPAPAKPMPRHEF YQPEEVGVALPAKGIKQNVNVEFGDQILSVVIDVAGEAVHFOPELFGK11PEKCR YEVLSUKLEAKAEIVTWASLEIGKQALLPKPNVASVRSORPVYPSKPKQKDWK LEAEVKQEKDEKLDGDAAMNKFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWEV GTKVVESTPPDGMELKKWEY"	
ORIGIN	Query March	Best Local Similarity	38.1%; Score 413.8; DB 8; Length 1197;	
			Matches 687; Conservative 1; Mismatches 373; Indels 20; Gaps 3;	
QY	7	CGCTCGGATCTGGAGAGCAGCCAGGAGCGCTTCGTGACGACGACGCTTCGAGCTGGCC	66	
DB	4	CGGAGCGNACTAGCGGAAAGCTAAGAGCGCTTCTAGATGACGACTTCGATGGCT	63	
QY	67	ACCGAGCTCTACAGCGAGCGCATCGAGCGCGGCGCGCCACCGCGACCTCTATGCGCAC	126	
DB	64	GCTGACTTATCTCAGAGCCATGACTTGGATGCTCTTTCGCGCTCTCTTCGCGCAT	123	
QY	127	CGCCCGGAGCGACATCAAGCTCGGCACTACACTGAGCGCTGGCGGATGCTTAAACAA	186	
DB	124	CGTGTCAAGCCACATCAACTCTTAACTTCAAGAGCTGTTCAGATGCGCAAAA	183	
QY	187	GCAATTGAGCTTGATCTCTATGATGATCAAAAGCTTACTACCGGAAAGGTGCTGCAATG	246	
DB	184	GCCATGAGTTGGAGCTTACTTTGGCCAAAGCTTATCTCAGAAAGGCGCTGCTTGATG	243	
QY	247	AAGCTTGAAGATACCAAACTGCAAGGCTGCTTGTAGTTGGTTCCTTCTTATGATCA	306	
DB	244	AAGCTAGAAGATATGCTTACTTAAAGCAGCTCTTGAGAGGCGCTTCGTTGCAACA	303	
QY	307	GGCGATTCAAGGTTGCTCGTCTATTGAAGNATGTGATGCGCATCGCTGAGGATCT	366	
DB	304	AATGAATCAAGTTTGAAGATGATGATGAATGCAACCTTCTCATTTGCAAGAA	360	
QY	367	AGCCAGGACCAAGTAAAGATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAGGAG	426	
DB	361	-----GAGAAAGATTGGTTCAACAGGTGCGCCAGCTTTGCTTCAAGTCT	407	
QY	427	GATTTCAAAATATGGAGATA-CAACACAGTGTAGTAACCCCAAGCAACCAATA	485	
DB	408	TACAACACCACTAGCAATCGAGCTGATGCTCTCTGCGACACCTGCGCAACCCATGTT	467	
QY	486	CAGGATCACTACTACACAGTGCCACAGAGAGTGTACTGCAATATTGCTAAGGGTGT	545	
DB	468	CAGACAGATTCTACCAAGAGCGGAGAGAGTGGGGGTGGCAATTTTCGCTAAGGGAT	527	
QY	546	TCCTGCTGATGTAGTCAITGATTTTGGTGAACAGATGTTGAGTGATCCATTTGAGT	605	
DB	528	CCCAAGCAGACGATGATGATGATGTTGGTGACAGATTTCTGAGTGTGATGATGAT	587	
QY	606	TCCTGCTGAGAACCATACCATTTTCAGCCCGCTGTTTCTAGATATCCCTGAGAA	665	

DB	588	TGCTGGAGAGAGCTTATCATTTTCAGCCAGATTGTTGGAGAGATATACCAGAGAA	647	
QY	666	ATGCAATATCAAGTCTTATCCACCAAGGTTGAAATACGCTTTCGCAAAAGCTGAGCAGT	725	
DB	648	GTGCAGATATGAGTATTTGTCGACCAAGTTTGAATCCGCTCTCGCAAAAGCAGAGATAGT	707	
QY	726	GACATGGACAAACCTGGATTTATAGTGGAGAGCAAGATCTGTTCCCAAGAGATAGCAC	785	
DB	708	CACTTGGGCTCCCTTGAATA---TGGCAAGGGCAAGCTTCTGCTTAAACCCCAATGT	764	
QY	786	GCCAGCTGAAACAGCCCAAGACCTTATACCCATCTTCAAGGCGGCAAAAGAGACTGGGA	845	
DB	765	TGCATCAGCAGTGTCTCAGAGGCGCGGTACCCATCTCTAAGCCCGGAAAGACTGGGA	824	
QY	846	TAACTGGAGCTGAAGTCAAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	905	
DB	825	CAAGCTGGAGCTGAAGTGAAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	884	
QY	906	ATTGAACAAATTTCTCCGTGACATCTAAGAGATGCTGATGAGATATTCGGAGGCGCAT	965	
DB	885	TATGAACAAATTTTTCAGGCGACATATACAGAGTGCAGATGAGGACATGAGGCGAGCAT	944	
QY	966	GATGAGTCACTGCTGAATCAATGCGACTGTTCTCAACCAATTGGAAGATGTTGG	1025	
DB	945	GAAACAAATTTTCAGAGTCCATGGAGCGTGTCTCTCAAACTGGAAGAGGTTGG	1004	
QY	1026	AGCAAGAGAGTGAAGAGGAGGCGCCCTGATGATGAGTATGAGCTCAAGAGTGGGAATCTA	1085	
DB	1005	GACAAAGAGTGGAGAGGACTCCACCAGATGCGATGGAGCTSAAGAGTGGGAGTATTG	1064	
QY	1086	A 1086		
DB	1065	A 1065		
RESULT 13				
AJ620882	Brassica oleracea mRNA for SGT1-like protein (sgtlb-1 gene).	1068 bp	mRNA	linear
LOCUS	Brassica oleracea mRNA for SGT1-like protein (sgtlb-1 gene).	1068 bp	mRNA	linear
DEFINITION	Brassica oleracea mRNA for SGT1-like protein (sgtlb-1 gene).	1068 bp	mRNA	linear
ACCESSION	AJ620882	1	GI:40974914	
VERSION	AJ620882.1	1	GI:40974914	
KEYWORDS	SGT1-like protein; sgtlb-1 gene.			
SOURCE	Brassica oleracea			
ORGANISM	Brassica oleracea			
REFERENCE	1	Yemm, A.A., Holub, E. and Tor, M.		
AUTHORS	Unpublished			
JOURNAL	2	(bases 1 to 1068)		
REFERENCE	Yemm, A.A.			
AUTHORS	Direct Submission			
TITLE	Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research,			
JOURNAL	Horticultural Research International, Wellesbourne, Warwick, UNITED			
FEATURES	KINGDOM	Location/Qualifiers		
source	1..1068	/organism="Brassica oleracea"		
	1..1068	/mol_type="mRNA"		
	1..1068	/variety="Senna"		
	1..1068	/db_xref="taxon:3712"		
	1..1068	/tissue_type="leaf"		
gene	1..1068	/gene="sgtlb-1"		
CDS	1..1068	/gene="sgtlb-1"		
	1..1068	/functions="role in disease resistance"		
	1..1068	/codon_start=1		
	1..1068	/evidence="experimental"		
	1..1068	/product="SGT1-like protein"		
	1..1068	/protein_id="CAF06580.1"		
	1..1068	/db_xref="GI:40974915"		

1054. .1084
/gene="At4g23570"

3' UTR

ORIGIN

Query Match	36.2%	Score 393.2;	DB 8;	Length 1084;
Best Local Similarity	62.7%	Pred. NO. 1.5e-60;		
Matches 677; Conservative	0;	Mismatches 373;	Indels 30;	Gaps 3;

QY	7	CGCTGGAATCTGGAGAGCAAGGCAAGGCGCTTCGTGACGACGACTCTTCGAGCTGCC	66
DB	4	GCGAAGAGGCTTGTGTAAGAGCTTAAGAGGCTTTCGTAGACGATGACTTCGATGTTGCT	63
QY	67	ACCGAGCTCTACAGCCAGGCGCATCGACGCGGGCCGCCACCGCGGACCTCTATGCCGAC	126
DB	64	GTTTGACTTGTATCTCAAAGCCATTGACTTGGATCCTTAATTCGCTGAGTCTTCGCTGAT	123
QY	127	CGCGCCGAGCGGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTTAACAAA	186
DB	124	CGTGCTCAGGCCATATCAAACTCGAAAGCTTCACTGAGGCGGTGGCAGATGCGAACAAA	183
QY	187	GCAATTGAGCTTGATCTCTATGATGCAATAAGCTTACTACGGAAAGGTGCTGATGATTT	246
DB	184	GCAATTGAGTTGGATCTCTCATTTGACCAAAAGCTTACCTAAGAAAGAACTGCTGTATG	243
QY	247	AAGCTTGAAGAAATACCAAACTGCAAAAGCTGCTCTGAGTTGGGTCTTCTTATGATCA	306
DB	244	AAGCTTGAAGATATCGGACTGCTTAACAAAGCTCTTGAAGAGGTGCTCTATCACGCCG	303
QY	307	GGCGAATCAAGGTTTGTCTGCTATTGAAGGAATGTGATAGCCGCACTCGTGAGGAATCT	366
DB	304	AGTGAATCCAAATTTAAGAGAGTTGATAGATGAATGCAATTTCTTAATCACAGAAGAA	360
QY	367	AGCCAGCACGATGAAGAAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG	426
DB	361	-----GAGAAGAATTTGGTTTCAACCGGTGCTTTCGACTTTTGCCTTCAAGTGT	407
QY	427	GATTTTCAAAATATGGAGAAATACACCAACAGTGTAGTAAACCCCAAGCAACCAAAATAC	486
DB	408	GACAGCACCA-----CCAGTATCTGAATCTGATGTCAACCCCTACAGCCNAGTAC	456
QY	487	AGGCATGACTACTACAAAGTGGCAGAGAGTGGTACTGCAATATTTGCTAAGGGTGTT	546
DB	457	AGGCACAGTACTACCAAAAGCCAGAGGAAGTTGTGTGTAAGTCTGTTTTCGAAAGGAATA	516
QY	547	CCTGCTGATAGTGTAGTCAATGATTTTGGTGGACAGATGTTTGATGTATCCATTGAAGTT	606
DB	517	CCCAAGCAGAAATGTTAAATCATCGACTTTGGTGAACAAATTCGTAGTGTGTGATTTGAGGTT	576
QY	607	CCTGGTGAAGAACCATACCAATTTTCAGCCGCTGTTTCTAAGATTTATCCCTTGAGAAA	666
DB	577	CCTGGAGAGATGCGTATTAATCTCCAACCGAGATGTTTGGAAAGATAATACCGATTAAG	636
QY	667	TGCAAAATATCAAGTCTTATCCCAACCAAGTTGAAATACGCCCTTGCAAAAGCTGAGCAGGTG	726
DB	637	TGCAAAATATGAAGTATTTGTGCGACAAAATTTGAGATCTGTCTTGCAAAAGCCGATATAATC	696
QY	727	ACATGGACAACCTCGGATTAATGTTGGGAAGACCAAAAGACTGTTTCCCAAGAGATGAAGCAG	786
DB	697	ACATGGGCTCTCCCTCGAAACCGGCAAAAGGGCCAGCGGTTTTGCAAAAGCCCTTAATGTCTCA	756
QY	787	CCAGCTGAAACAGCCCCCAAGACCTTCATACCCATCTTCAAGGCGGAAAAAGACTCGGAT	846
DB	757	TCAGAGGTTTCA---CAGGGCCAGCTTATTCATCTTCTTAAGAAAGTTAAGGCTGGGAC	813
QY	847	AAACTGGAAGCTGAAGTCAAAAAGGAGGAGGAAGGAAAAAATTTGATGGTGTGCTGCA	906
DB	814	AAGCTGAAGCTGAAGTGAAGACGACGAGAAAAAGATGAGAAGCTGGAAGGAGAGCGCAGCT	873
QY	907	TTGAAACAAATTTCTTCGTGACATCTACAAGATGCTGATGAAGATATGCGGAGGGCCCATG	966
DB	874	TTGAACAAATTTCTTCGCGGAGATATACAGAAATGCGGATGAGATATGAGCGCAGCAGT	933

Qy	967	ATGAGTCAATCTGTGGAATCAANTGCGACTGTGTCTCTCAACCCAATTGGAAAGATGTTGA	1026
Db	934	AGCAAAATCATTTGTGGAATCGAAATGGGACAGTCTGTCAAACAATCGGCAAGGTTGGG	993
Qy	1027	GCAAGAAGGTAGAAAGGAGCCCCCTGTATGTTGATCGAGCTCAAGAAGTGGGAATACTAA	1086
Db	994	ACTTAAGACAATCGAGAGTACTCTCCGGATGGCATTGGAGCTCAAGAAATGGGAGATCTGA	1053

Search completed: April 23, 2005, 08:51:29
Job time : 4822.69 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 00:00:36 ; Search time 3903.36 Seconds
(without alignments)
10590.311 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggcgcgtcgatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_ges1.*

9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1088.4	98.4	1646	3	AY103953	Zea mays
2	698.2	64.3	1131	9	CU959360	OSIFCC002
3	652.4	60.1	668	2	AW076274	614063D11
4	650.2	59.9	752	6	CD443320	EL01N0424
5	647	59.6	756	6	CA184150	SCRFST314
6	643.4	59.2	659	2	AW090936	614068G07
7	643.4	59.2	772	6	CA258233	SCCCT3300
8	641	59.0	759	6	CA245820	SCSZFL508
9	631	58.1	666	6	CA215140	SCSBAD112
10	608.2	56.0	678	6	CA177070	SCMST1105
11	607	55.9	665	6	CA153735	SCVPR2203
12	596.4	54.9	622	6	CA223980	SCJFFL3C0
13	583.4	53.7	586	2	AW042392	61402BB03
14	583.4	53.7	609	6	CA229238	SCAGFL302
15	572	52.7	702	6	CA262995	SCPRLB202
16	566	52.1	718	6	CA106435	SCQCHR101
17	565.2	52.0	594	6	CA098274	SCMCC1605
18	562	51.7	574	4	BI644145	949022C08
19	557.2	51.3	630	6	CA486265	WHE4329 D
20	555.6	51.2	579	2	AW066518	660015G08
21	552.8	50.9	680	6	CA128103	SCAGLR203
22	550	50.6	550	4	BI679413	949001E02
23	550	50.6	914	6	CA084222	SCSQAW203
24	545	50.2	588	5	BU499669	94617BB11

25	541	49.8	601	6	CA176126	CA176126 SCJLST102
26	535.8	49.3	553	5	BU098322	946134B07
27	529.4	48.7	612	6	CA234143	SCCCCL6C0
28	523	48.2	615	6	CA283412	SCSBS105
29	522.6	48.1	730	2	BE362001	DGL_83_H0
30	520.8	48.0	725	6	CA094787	SCCCL4701
31	520	47.9	578	6	CD996630	QBC6a07_X
32	519.6	47.8	708	6	CA240596	SCSBL406
33	514.8	47.4	1089	7	CK208966	FGAS02069
34	512.2	47.2	693	6	CA104899	SCJPHR1C0
35	511.6	47.1	606	6	CA083869	SCBPAM205
36	508.4	46.8	625	6	CA266503	SCAGLB204
37	507.4	46.7	518	2	AW057009	660007D01
38	507.4	46.7	862	6	CB655564	OSJNEC09B
39	507.2	46.7	652	6	CA102017	SCAGHR101
40	505.2	46.5	798	6	CA176210	SCJLST102
41	500.6	46.1	610	6	CA184034	SCQST311
42	499.4	46.0	643	6	CA251380	SCQSF112
43	498.8	45.9	778	6	CA227695	SCJLFL301
44	498.4	45.9	607	6	CA120629	SCCCLR108
45	487.4	44.9	792	6	CB657723	OSJNEC13E

ALIGNMENTS

RESULT 1	AY103953	1646 bp	linear	HTC	16-OCT-2002
LOCUS	AY103953	Zea mays	PCO098564	mRNA sequence.	
DEFINITION	AY103953	HTC.			
ACCESSION	AY103953.1	GI:21207031			
VERSION	HTC.				
KEYWORDS	Zea mays				
SOURCE	Zea mays				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
REFERENCE	1 (bases 1 to 1646)				
AUTHORS	Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 1646)				
AUTHORS	Coe, E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				

FEATURES	Location/Qualifiers
source	1..1646
	/organism="Zea mays"
	/mol_type="mRNA"
	/db_xref="MaizeDB:635643"
	/db_xref="taxon:4577"
	/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
	/notes="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
ORIGIN	
Query Match	98.4%; Score 1068.4; DB 3; Length 1646;

364	Qy	TCTAGCCAGGCACACAGTAAAGAAATGTTGAGG-----CTACTGTGGCTGCT	408
373	Db	CTTAGTGAAGTCCCTGTGTTAAGAAAGGCTGAAGATGGAGCAGCTGGCCCTCTGTGTTGCTTCT	432
409	Qy	ACTATTGAGGCAAGGAGGAGTATTCACAAATATGGAGAATACACACCAGTCATAGAACCC	468
433	Db	TTTGTTGAGGAAAGAGATGATGCTGCAAAATATGATATATACACCCAAATGGTAGA-----	488
469	Qy	CCAAGCAAAACCAAAATACAGGCATGACTACTACAAACAGTCCACAGAAAGTGTTACTGACA	528
489	Db	--AGTGAAGCCAAAATAACAGGCACGACTTCTACAAACAGTGCTACAGAAAGTTGTTATGACA	546
529	Qy	ATATTTCTGAAGGCTGTTCTCGTCGTATAGTGTAGTCAATTGATTTTGGTGAAACAGATGTTG	588
547	Db	ATTTTTGCCAAAGGGGTGTTCTCTGCTGAGAAATGTTGTTGTGATTTTGGTGAAACAAATGTTA	606
589	Qy	AGTGATATCCATTGAAGTTCTCTGGTGAAAGAACCATACCACTTTTCAGCCCCGTCGTGTTTCT	648
607	Db	AGTGTGTGATTTGAAGTCCCTTGGAGAGGAGCCGTACCAATTTTCAGCTCTGCTGTTTCT	666
649	Qy	AAGATTATCCCTTGAGAAATGCAAAATATCAAGTCTTATTCACCAAGGTTGAAATACGCCTT	708
667	Db	AAGATCATCTCTGAGAAAAGCAGATACCAAGTGTCTATCCACGAAGTTGAAATTAAGACTG	726
709	Qy	GCAAAAGCTGAGCAGGTTGACATGAGCAACCCCTGGAATTATAGTGAAGACCAAGACTGTT	768
727	Db	GCTAAAGCTGAACAGATTATCATGACCTCACTTGATTATGATAAAAAACCAAGGCTGTT	786
769	Qy	CCCAGAGAAGATAAGCAGC-----CAGCTGAAACAGCC	801
787	Db	CCACAAAGATATCCCTCCAGTTTACTGATAGCCCTCTTTTATCAGCTGAAATCGGCC	846
802	Qy	CCAAGACTTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAATCGAAGCTGAA	861
847	Db	CAGAGGCCATCATATCTCTCTCAAATCCAAAGAAAGACTGGGATAAATCGAAGCTGAA	906
862	Qy	GTCAAAAGAGGAGGAAGGAAGAAAATCTTGATGGTGATGCTGATTGAAACAAATCTTTC	921
907	Db	GTTAAAAAGGAGGAGAGGAGGAGAAGCTTGAAGCGATGCTGCAATTGAAACAAATTTTTC	966
922	Qy	CGTGACATCTCAAGGATGCTGATGAAGATATCGGAGGGCCATGATCAAGTCAATCTCGT	981
967	Db	CGTGACATCTACGTGATGCTGATGAAGACATGCGACGAGCAATGATGAATCTTTTGT	1026
982	Qy	GAATCAAATGGCACTGTTTCTCTCAACCAATGGAAAGATGTTGGAGCAAAAGAGTGAA	1041
1027	Db	GAATCTAACGGTACTGTTCTGTGACCAATTTGGAAGAGATGTTGCTCGAAGAAAGTAGAG	1086
1042	Qy	GGGAGCCCCCTGATGTATGGAGCTCAAGAAAGTGGGAATCTAA	1086
1087	Db	GGAAGCCCACTGATGGGATGGAGCTTAAGAAATGGGAGTACTAA	1131

RESULT 3	AW076274	AW076274	668 bp	mRNA	linear	EST 14-OCT-1999
LOCUS						
DEFINITION						

KEYWORDS	EST.
SOURCE	Zea mays
ORGANISM	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 668)

AUTHORS	walbot.v.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	University
COMMENT	Unpublished (1999) Contact: Walbot V Department of Biological Sciences

```

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614063 row: D column: 11.
Location/Qualifiers
1. .668
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db xref="taxon:4577"
/tissue_type="root"
/dev stage="3-4 days old"
/lab_host="XLOLR"
/clone lib="G14 - root cDNA library from Walbot Lab"
/notes="Organ: root; Vector: pBlueScriptII SK+; Site 1:
ECORI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
FEATURES
source

```

ORIGIN

Query Match	60.1%;	Score 652.4;	DB 2;	Length 668;
Best Local Similarity	99.8%;	Pred. No. 1.le-171;		
Matches	653;	Conservative	0;	Mismatches 1;
			Indels	0;
			Gaps	0;
Qy	397	ACTGTGGCTCTACTATTGTGAGGCAAGGAGGATTTTCA	CAAAATATGGAGNATACACCACCA	456
Db	14	ACTGTGGCTCTACTATTGTGAGGCAAGGAGGATTTTCA	AAATATGGAGNATACACCACCA	73
Qy	457	GTGATAGAACCCCAAGCAAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA		516
Db	74	GTGATAGAACCCCAAGCAAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA		133
Qy	517	GTGGTACTGACAATATTTGCTAAAGGGTGTTCCTGCTGATAGTGTAGTCATTTGGT		576
Db	134	GTGGTACTGACAATATTTGCTAAAGGGTGTTCCTGCTGATAGTGTAGTCATTTGGT		193
Qy	577	GAACAGATGTTGATGTATCCATTTGAAGTTCTCTGGTGAGAAACCATATTTTCAGCCC		636
Db	194	GAACAGATGTTGATGTATCCATTTGAAGTTCTCTGGTGAGAAACCATATTTTCAGCCC		253
Qy	637	CGTCTGTTTTCTAAGATTTATCCCTTGAGAAATGCAAAATCAAGTCTTATCCACCAAGGTT		696
Db	254	CGTCTGTTTTCTAAGATTTATCCCTTGAGAAATGCAAAATCAAGTCTTATCCACCAAGGTC		313
Qy	697	GAATACGCTTGC AAAAGCTGACGAGGTGACATGGACAACCCCTGGATTTATAGTGGAGA		756
Db	314	GAATACGCTTGC AAAAGCTGACGAGGTGACATGGACAACCCCTGGATTTATAGTGGAGA		373
Qy	757	CCAAAGACTGTTCCCAAGAGATGAAGCACGCCAGCTGAAAACAGCCCCCAAGACCTTTCATAC		816
Db	374	CCAAAGACTGTTCCCAAGAGATGAAGCACGCCAGCTGAAAACAGCCCCCAAGACCTTTCATAC		433
Qy	817	CGATCTTCAAAGCGGAAAAAAGACTGGGATAAATCGGAAGCTGAAGTCAAAAAGAGGAG		876
Db	434	CCA TCTTCAAAGCGGAAAAAAGACTGGGATAAATCGGAAGCTGAAGTCAAAAAGAGGAG		493
Qy	877	AAGGAAGAAAACTTGATGTGTGATGCTGCAATGGAACAAAATTTCTTCCGTGACATCTACAAG		936
Db	494	AAGGAAGAAAACTTGATGTGTGATGCTGCAATGGAACAAAATTTCTTCCGTGACATCTACAAG		553
Qy	937	GATGCTGATGAAGATATCGGAGGGCCCATGATGAAGTCATTTCGTGGAAATCAAATGGCACT		996
Db	554	GATGCTGATGAAGATATCGGAGGGCCCATGATGAAGTCATTTCGTGGAAATCAAATGGCACT		613
Qy	997	GTTCTCTCAACCAATTTGGAAAGATGTTTGGAGCAAAAGAAAGGTAGAAAGGAGCCCC		1050
Db	614	GTTCTCTCAACCAATTTGGAAAGATGTTTGGAGCAAAAGAAAGGTAGAAAGGAGCCCC		667

RESULT 4
CD443320
LOCUS

Contact: walbot.v.
Department of Biological Sciences

752 bp mRNA linear EST 03-JUN-2003

CD443320

```

DEFINITION EL01N0424G10.b Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION CD443320
VERSION CD443320.1 GI:31358963
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 752)
AUTHORS Lai, J., Dey, N., Kim, C.S., Bharti, A.K., Rudd, S., Mayer, K.F.,
Larkins, B., Becraft, P. and Messing, J.
TITLE Characterization of the maize endosperm transcriptome and its
comparison to the rice genome
JOURNAL Comp. Res. 14 (10), 1932-1937 (2004)
COMMENT Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
FEATURES             Location/Qualifiers
     source           1..752
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W22"
                     /db_xref="taxon:4577"
                     /tissue_type="Endosperm of 7-23DAP"
                     /clone_lib="Endosperm 4"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"

ORIGIN
Query Match      59.9%; Score 650.2; DB 6; Length 752;
Best Local Similarity 99.5%; Pred. No. 4.9e-171;
Matches 652; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCCGCGCTGGATCTGGAGGAGCAAGCCCAAGGAGGCTTCTGTCGACGACGACTTCGAG 60
DB 98 ATGCCGCGCTGGATCTGGAGGAGCAAGCCCAAGGAGGCTTCTGTCGACGACGACTTCGAG 157

QY 61 CTGGCCACCGAGCTCTACAGCCAGGCGCCATCGACGCGCGCGCCGCGCACCGCGGACTCTAT 120
DB 158 CTGGCCACCGAGCTCTACAGCCAGGCGCCATCGAGCGCGCGCGCCGCGCACCGCGGACTCTAT 217

QY 121 GCCGACCGCGCGCGCGCGCACATCAAGCTCGGCACTACACTGAGGCTGTGGCGGATGCT 180
DB 218 GCCGACCGCGCGCGCGCGCACATCAAGCTCGGCACTACACTGAGGCTGTGGCGGATGCT 277

QY 181 AACAAAGCAATTGAGCTTGATCTTATGATGATTAAGCTTACTACCGGAAAGGTGCTGCA 240
DB 278 AACAAAGCAATTGAGCTTGATCTTATGATGATTAAGCTTACTACCGGAAAGGTGCTGCA 337

QY 241 TGCATTAAAGTTGAAGATACAAATCGCAAGGCTGCTTGTAGTGGGTTCTTCTTAT 300
DB 338 TGCATTAAAGTTGAAGATACAAATCGCAAGGCTGCTTGTAGTGGGTTCTTCTTAT 397

QY 301 GCATCAGCGGATCAAGGTTTCTGCTTATTTGAAGGATGTGATGAGCGATCGCTGAG 360
DB 398 GCATCAGCGGATCAAGGTTTCTGCTTATTTGAAGGATGTGATGAGCGATCGCTGAG 457

QY 361 GAATCTAGCCAGGACCAAGTAAGATTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
DB 458 GAATCTAGCCAGGACCAAGTAAGATTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 517

QY 421 AAGGAGGATTTCAAAATATGAGAAATACACCAACAGTGATAGAACCCCGCAAGCAACCA 480
DB 518 AAGGAGGATTTCAAAATATGAGAAATACACCAACAGTGATAGAACCCCGCAAGCAACCA 577

QY 481 AATACAGGCATGACTACTACAAAGTGCCACAGAGTGGTGCTACTGACAAATTTTCTTAAG 540

```

```

578 AAATACAGCATGACTACTACAAAGTGGTACTGACAAATTTTGTCTAAG 637
541 GGTGTTCTCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTGAGTGTATCCATT 600
638 GGTGTTCTCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTGAGTGTATCCATT 697
601 GAAGTTCTCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTGAGTGTATCCATT 655
698 GAAGTTCTCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTGAGTGTATCCATT 752

RESULT 5
CA184150 756 bp mRNA linear EST 24-SEP-2003
LOCUS SCRFST3142H05.9 ST3 Saccharum officinarum cDNA clone SCRFST3142H05
5', mRNA sequence.
CA184150 GI:35121633
VERSION Saccharum officinarum
KEYWORDS Saccharum officinarum
SOURCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ORGANISM Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 756)
AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
TITLE The libraries that made SUCST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 142 row: H column: 05
Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source           1..756
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCRFST3142H05"
                     /lab_host="DH10B"
                     /clone_lib="ST3"
                     /note="Organ: Fourth apical stalk internodes of adult
                     plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Fourth apical
                     stalk internodes of adult plants]. cDNA was prepared from
                     polyA+ mRNA using SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were fractionated
                     in a sapharose CI-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details of each source of RNA and library
                     construction can be obtained at
                     http://sucst.lad.ic.unicamp.br/public"

ORIGIN
Query Match      59.6%; Score 647; DB 6; Length 756;
Best Local Similarity 95.0%; Pred. No. 3.9e-170;
Matches 721; Conservative 0; Mismatches 32; Indels 6; Gaps 5;

QY 4 GCCCGCTCGGATCTGGAGAGCAAGCCCAAGAGGCGCTTCGTCGACGACGACTTCGAGCTG 63
DB 3 GCCCGCTCGGATCTGGAGAGCAAGCCCAAGAGGCGCTTCGTCGACGACGACTTCGAGCTG 62
QY 64 GCCACCGAGTCTTACAGCCAGGCGCATCGACCGCGGCGCGCCACCGCGGCGCTTATGCC 123
DB 63 GCGCGCGAG--CTACACCCAGGCGCATCGACCGCGGCGCGCCACCGCGGCGCTTATGCC 120

```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.

1. (bases 1 to 772)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.

The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bcccenter.fcav.unesp.br>

Plate: 004 row: D column: 05

Seq primer: T7 Promoter Primer.

FEATURES

source

1. 772

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clones="SCCCT3004D05"

/lab_host="DH10B"

/clone_lib="RT3"

/note="Organ: Root apex from adult plants; Vector:

pSport1; Site_1: SalI; Site_2: NotI; An unidirectional

cDNA library generated from [Root apex from adult

plants]. cDNA was prepared from polyA+ mRNA using

SuperScript Plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a sepharose

CL-2B 40cm-columns and fragments sizing between 0.8 and

1.5 Kb were directionally cloned into the vector. Details

of each source of RNA and library construction can be

obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 59.2%; Score 643.4; DB 6; Length 772;
Best Local Similarity 96.9%; Pred. No. 4e-169;
Matches 656; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

4 GCCCGCTCGGATCTGGAGACGACCAAGCCAGGAGCGCTTCGTGACGACGACTTCGAGCTG 63
|||||
96 GCCCGCTCGGATCTGGAGACGACCAAGCCAGGAGCGCTTCGTGACGACGACTTCGAGCTG 155
|||||

64 GCCACCGAGCTCTAAGCCAGGCGCATCGACCGCGGCGCGCCGACCGCGGACCTCTATGCC 123
|||||

156 GCCGCGGAGCTCTACACCGAGGCGCATCGACCGCGGCGCGCGCGCGGACCTCTAGGCC 215
|||||

124 GACCGCGCGGCGGCGCATCGACCTCGGCGCACTACACTGAGGCTGTGGCGGATGCTAAC 183
|||||

216 GACCGCGCGGCGGCGCATCAAGCTCGGCGCACTACACTGAGGCTGTGGCGGATGCTAAC 275
|||||

184 AAAGCAATTGAGCTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 243
|||||

276 AAAGCAATTGAGCTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 335
|||||

244 ATTAAGCTTGAGGAATACAACTGCAAGGCTGCTCTTGATGTTGGGTTCTTCTATGCA 303
|||||

336 ATTAAGCTTGAGGAATACAACTGCAAGGCTGCTCTTGATGTTGGGTTCTTCTATGCA 395
|||||

304 TCAGGCGATTCAAGGTTTCTGCTCTATTGAGGAATGATGATGATGATGATGATGATGATG 363
|||||

396 TCAGGCGATTCAAGGTTTCTGCTCTATTGAGGAATGATGATGATGATGATGATGATGATG 455
|||||

364 TCTAGCCAGGACCAAGTAAGAAATGTTGAGGCTACTGTGGCTGCTACTATTGAGCAAG 423
|||||

456 TCTAGCCAGGACCAAGTAAGAAATGTTGAGGCTCTCTGTGGCTGCTACTATTGAGCAAG 515
|||||

424 GAGGATTTACAAATATGAGGATACACCAAGGATGATGATGATGATGATGATGATGATG 483
|||||

Db 516 GAGATGCCACAATATGGAGATATACCGCCAGTGGTAGACCCCAAGCAACCTTAA 575
|||||

Qy 484 TACAGGCATGACTACTACCAACAGTGCACAGAGTGGTACTGACAATATTTGTAAGGGT 543
|||||

Db 576 TACAGGCATGACTACTACCAACAGTGCACAGAGTGGTACTGACAATATTTGTAAGGGT 635
|||||

Qy 544 GTTCCTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTAGTGTATCCATTGAA 603
|||||

Db 636 GTTCCTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTAGTGTATCCATTGAA 695
|||||

Qy 604 GTTCCTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTAGTGTATCCATTGAG 663
|||||

Db 696 GTTCCTGCTGATAGTGTAGTCAATGATTTTGGTGAACAGATGTTAGTGTATCCATTGAG 755
|||||

Qy 664 AAATGCAAAATATCAAGT 680
|||||

Db 756 AAATGCAAAATATCAAGT 772
|||||

RESULT 8

CA245820

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA245820 759 bp mRNA linear EST 25-SEP-2003
SCEZFL5084D06.g Saccharum officinarum FL5 Saccharum officinarum
cDNA clone SCEZFL5084D06 5', mRNA sequence.

CA245820

CA245820.1 GI:35324833

EST.

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

1. (bases 1 to 759)

Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bcccenter.fcav.unesp.br>

Plate: 084 row: D column: 06

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. 759

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clones="SCEZFL5084D06"

/lab_host="DH10B"

/clone_lib="Saccharum officinarum FL5"

/note="Organ: Developed inflorescence (20cm-long) without

rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [Developed

inflorescence (20cm-long) without rachis]. cDNA was

prepared from polyA+ mRNA using SuperScript Plasmid

System Kit (Invitrogen). The double-strand cDNAs were

fractionated in a sepharose CL-2B 40cm-columns and

fragments sizing between 0.8 and 1.5 Kb were

directionally cloned into the vector. Details of each

source of RNA and library construction can be obtained at

<http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 59.0%; Score 641; DB 6; Length 759;
Best Local Similarity 93.6%; Pred. No. 1.9e-168;

Matches 700; Conservative 0; Mismatches 43; Indels 5; Gaps 3;

Qy 207 GATCATTAAGAGCTTACTACCGGAAGGTGCTGATGATTAAGCTTGAAGATACCAAC 265
Db 2 GATCATTAAGAGCTTACTACCGGAAGGTGCTGATGATTAAGCTTGAAGATACCAAC 61

Qy 267 TGCAAGGCTGCTCTGAGTTGGGTCTTCTTATGCAATCAGGCGATTCAGAGTTTGCTCG 326
Db 62 TGCAAGGCTGCTCTGAGTTGGGTCTTCTTATGCAATCAGGCGATTCAGAGTTTGCTCG 121

Qy 327 TCTATTGAAGGAATGTTGATGAGCGCATCGCTGAGGAATCTAGCCAGGACCAGTAAGAA 386
Db 122 TCTATTGAAGGAATGTTGATGAGCGCATCGCTGAGGAATCTAGCCAGGACCAGTAAGAA 181

Qy 387 TGTTGAGCTACTGTGCTGCTACTATTTAGGACAGAGGATTTCAAAATATGGAGAA 446
Db 182 GGTTCAGGCTCTGCTGCTGCTACTGTTGAGGACAGAGGATTCACAAATATGGAGAA 241

Qy 447 TACACCAAGCTGATAGAACCCCAAGCAACCAAAATACAGGATGACTACTACAACAG 506
Db 242 TACACCGCAGTGTAGAACCCCAAGCAACCAAAATACAGGATGACTACTACAACAG 301

Qy 507 TGCCACAGAAGTGTACTGACAAATATTTGCTAAGGGTGTCTCTGCTGATAGTGTAGTCAT 566
Db 302 TGCCACAGAAGTGTCTGCTGACAAATATTTGCTAAGGGTGTCTCTGCTGATAGTGTAGTCAT 361

Qy 567 TGATTTTGGTGAACAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
Db 362 TGATTTTGGTGAACAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 421

Qy 627 TTTTCAGCCCGCTGCTGTTTCTTAAGATTTATCCCTGAGAAATGCAAAATCAAGTCTTATC 686
Db 422 TTTTCAGCCCGCTGCTGTTTCTTAAGATTTATCCCTGAGAAATGCAAAATCAAGTCTTATC 481

Qy 687 CACCAAGGTTGAATACGCTTTCGAAAGCTGAGCAGTGAATGAGCAACCCCTGGATTA 746
Db 482 CACCAAGGTTGAATACGCTTTCGAAAGCTGAGCAGTGAATGAGCAACCCCTGGATTA 541

Qy 747 TAGTGAAGACCAAGAGTGTTCCTCCAGAGATTAAGCAGCGCAGTGAACAGCCCAAG 806
Db 542 TAGTGAAGACCAAGAGTGTTCCTCCAGAGATTAAGCAGCGCAGTGAACAGCCCAAG 601

Qy 807 ACCTTCATACCATCTTCAAGGCGGAAAGAGCTGG--GATAACTGGAAGCTGAAGT 863
Db 602 ACCTTCATACCATCTTCAAGGCGGAAAGAGAGCTGGGNGATAACTGGGAGCTGAGTC 661

Qy 864 CAAAAGGAGGAGAA--GGAAGAAAACCTTGATGTTGATGCTGCAATGAAACAAATTTCTCC 922
Db 662 AAAAGGAGGAGAGGAGGAGAAACTTGATGTTGATGCTGATGCTTGNNAACATTC--TCC 720

Qy 923 GTGACATCTACAAGATGCTGATGAAGA 950
Db 721 GTGACATCTACAAGATGCTGATGAAGA 748

RESULT 9

CA215140

LOCUS

DEFINITION SCSBAD1128H12.g AD1 Saccharum officinarum cDNA clone SCSBAD1128H12 5', mRNA sequence.

ACCESSION

CA215140

VERSION

CA215140.1 GI:35263122

KEYWORDS

EST.

SOURCE

Saccharum officinarum

ORGANISM

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

1 (bases 1 to 666)

Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

JOURNAL

COMMENT

Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccccenter.fcav.unesp.br>
Plate: 128 row: H column: 12
Seq primer: T7 Promoter Primer.

FEATURES

source

1..666
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCSBAD1128H12"
/lab_host="DH10B"
/clone_1lb="AD1"
/notes="Organ: seedlings inoculated with Gluconacetobacter diazotrophicans; Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [seedlings inoculated with Gluconacetobacter diazotrophicans]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 58.1%; Score 631; DB 6; Length 666;
Best Local Similarity 96.5%; Pred. No. 1.2e-165;
Matches 643; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 339 ATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGACCAAGATGTTGAGGCTAC 398
Db 1 ATGTGATGAGTGCATCGCTGAGGAATCTAGCCAGGACCAAGATGTTGAGGCTCC 60

Qy 399 TGTGGCTGTCTACTATTGAGGACCAAGGAGATTTCAAAATATGGAGATACCAACCAAGT 458
Db 61 TGTGGCTGTCTACTATTGAGGACCAAGGAGATGCAAAATATGGAGATACCAACCAAGT 120

Qy 459 GATAGAACCCCAAGCAACCAAAATACAGGATGACTACTACAAAGTGCACAGAGT 518
Db 121 GGTAGAACCCCAAGCAACCAAAATACAGGATGACTACTACAAAGTGCACAGAGT 180

Qy 519 GGTACTGACAAATATTTGCTAAGGGTGTCTCTGCTGATGATGATGATGATGATGATGAT 578
Db 181 GGTGCTGACAAATATTTGCTAAGGGTGTCTCTGCTGATGATGATGATGATGATGATGAT 240

Qy 579 ACAGATGTTGAGTGTATCCATTGAAGTTCCTGTTGGAAGAACCAATACCAATTTTACGCCCG 638
Db 241 ACAGATGTTGAGTGTATCCATTGAAGTTCCTGTTGGAAGAACCAATACCAATTTTACGCCCG 300

Qy 639 TCTGTTTTCTAAGATATCCCTGAGAAATCAAAATATCAAGTCTTATCCAAAGGTTGA 698
Db 301 TCTGTTTTCTAAGATATCCCTGAGAAATCAAAATATCAAGTCTTATCCAAAGGTTGA 360

Qy 699 AATACGCTTTGCAAAAGCTGAGCGATGACATGGACCAACCTTGGATTTATAGTGAAGACC 758
Db 361 AATACGCTTTGCAAAAGCTGAGCGATGACATGGACCAACCTTGGATTTATAGTGAAGACC 420

Qy 759 AAGAGCTGTTTCCCAAGAGATAGCAGCCAGCTGAAACAGCCCAAGACCTTCATACCC 818
Db 421 AAGAGCTGTTTCCCAAGAGATAGCAGCCAGCTGAAACAGCCCAAGACCTTCATACCC 480

Qy 819 ATCTTCAAGCGCAAAAGAGCTGGGATAAATCTGGAAGCTGAACTCAAAAGGAGAGAA 878
Db 481 ATCTTCAAGCGCAAAAGAGAGCTGGGATAAATCTGGAAGCTGAACTCAAAAGGAGAGAA 540


```

QY 879 GGAAGAAAACCTTGATGCTGCAATGAAACAAATCTTCCGTGACATCTACAAGGA 938
Db |||||
541 GGAAGAAAACCTTGATGCTGCAATGAAACAAATCTTCCGTGACATCTACAAGGA 600
QY 939 TCCTCATGAAGATATCCGAGGGCCATGATGAAGTCATTCGTGGAATCAATGGCAGCTGT 998
Db |||||
601 TCCTCATGAAGATATCCGAGGGCCATGATGAAGTCATTCGTGGAATCAATGGCAGCTGT 660
QY 999 TCTCTC 1004
Db |||||
661 TCTCTC 666

RESULT 10
CA177070
LOCUS SCMCST1057G04.g ST1 Saccharum officinarum cDNA clone SCMCST1057G04
DEFINITION 5', mRNA sequence.
ACCESSION CA177070
VERSION CA177070.1 GI:35107470
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 678)
Vettore A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 057 row: G column: 04
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1 .678
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCMCST1057G04"
/lab_host="DH10B"
/clone_lib="ST1"
/note="Organ: First apical stalk internodes of adult
plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [First apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

FEATURES
source
1 .678
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCMCST1057G04"
/lab_host="DH10B"
/clone_lib="ST1"
/note="Organ: First apical stalk internodes of adult
plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [First apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Query Match 56.0%; Score 608.2; DB 6; Length 678;
Best Local Similarity 95.0%; Pred. No. 3e-159;
Matches 628; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 208 ATGCATAAAGCTTACTACCGAAAGGTCCTGCATGCAATTAAGCTTGAAGAATACCAACT 267
Db |||||
18 ACGCGTCGGGCTACTACCGAAAGGTCCTGCATGCAATTAAGCTTGAAGAATACCAACT 77
QY 268 GCAAAGGCTGCTCTTGAGTTGGGTCTCTTATGCATCAGGCGATTCAGGTTTGCTCGT 327

```

```

Db |||||
78 GCAAAGGCTGCTCTTGAGTTGGGTCTCTTATGCATCAGGCGATTCAGGTTTACTCGT 137
QY 328 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCAACGATTAAGAAT 387
Db |||||
138 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCAACGATTAAGAAT 197
QY 388 GTTCAGGCTACTGCTGGCTGCTACTATTGAGGACAAGGAGGATTTTCACAAATATCGAAT 447
Db |||||
198 GTTCAGGCTCTGCTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATCGAAT 257
QY 448 ACACCCAGGTGATAGAACCCCAAGCAAAACCAAAATACAGGCATGACTACTACCAAGT 507
Db |||||
258 ACACCCAGGTGATAGAACCCCAAGCAAAACCAAAATACAGGCATGACTACTACCAAGT 317
QY 508 GCACAGAGTGTGATGACAAATATTTGCTAAGGGTGTCTGCTGATAGTGTAGTCAAT 567
Db |||||
318 GCACAGAGTGTGCTGACAAATATTTGCTAAGGGTGTCTGCTGATAGTGTAGTCAAT 377
QY 568 GATTTGCTGAACAGATGTTGAGTGTATCCATTGAAGTCTCTGCTGAGAACCATCAT 627
Db |||||
378 GATTTGCTGAACAGATGTTGAGTGTATCCATTGAAGTCTCTGCTGAGAACCATCAT 437
QY 628 TTTGAGCCCGCTCTGTTTCTTAAGATTTATCCCTGAGAAATGCAAAATCAAGTCTTATCC 687
Db |||||
438 TTTGAGCCCGCTCTGTTTCTTAAGATTTATCCCTGAGAAATGCAAAATCAAGTCTTATCC 497
QY 688 ACCAGGTTGAATACCGCTTGCAAAAGCTGAGCAGTGAACAGCCCTGGATAT 747
Db |||||
498 ACCAGGTCGAAATACCGCTTGCAAAAGCTGAGCAGTGAACAGCCCTGGATAT 557
QY 748 AGTGAAGACCAAGACTGTTCCCAAGATAAGCAGCAGCTGAACAGCCCAAGA 807
Db |||||
558 AGTGAAGACCAAGACTGTTCCCAAGATAAGCAGCAGCTGAACAGCCCAAGA 617
QY 808 CTTTCATACCCATCTTCAAAAGGCGAAAGAGCTGGGATAAAGCTGGAAGTGAAGTCAAA 867
Db |||||
618 CTTTCATACCCATCTTCAAAAGGCGAAAGAGCTGGGATAAAGCTGGAAGTGAAGTCAAA 677
QY 868 A 868
Db 678 A 678

RESULT 11
CA153735
LOCUS SCVPRZ2035E06.g RZ2 Saccharum officinarum cDNA clone SCVPRZ2035E06
DEFINITION 5', mRNA sequence.
ACCESSION CA153735
VERSION CA153735.1 GI:35060111
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 665)
Vettore A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 035 row: E column: 06

```


Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..665
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCVPR22035E06"
/lab_host="DH10B"
/clone_lib="R22"
/note="Organ: Shoot-root transition zone from young plants (small insert library); Vector: pSport1; Site: 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [shoot-root transition zone from young plants (small insert library)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

FEATURES
source
1..622
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFL3C02C11"
/lab_host="DH10B"
/clone_lib="Saccharum officinarum FL3"
/note="Organ: Base of developing inflorescence (5cm-long); Vector: pSport1; Site: 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [base of developing inflorescence (5cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

LOCUS
CA229980 622 bp mRNA linear EST 25-SEP-2003
DEFINITION
SCJFL3C02C11.g Saccharum officinarum FL3 Saccharum officinarum cDNA clone SCJFL3C02C11 5', mRNA sequence.
ACCESSION
CA229980
VERSION
CA229980.1 GI:35293036
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE
1 (bases 1 to 622)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccc.net.br>
Plate: C02 row: C column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..622
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFL3C02C11"
/lab_host="DH10B"
/clone_lib="Saccharum officinarum FL3"
/note="Organ: Base of developing inflorescence (5cm-long); Vector: pSport1; Site: 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [base of developing inflorescence (5cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN
Query Match 54.9%; Score 596.4; DB 6; Length 622;
Best Local Similarity 97.4%; Pred. No. 5.9e-156;
Matches 606; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 262 CAAACTCAAGGCTGCTTTGAGTGGTCTTTCTTATGTCATCAGGCGATTCAAGGTTT 321
DB 1 CAAACTCAAGGCTGCTTTGAGTGGTCTTTCTTATGTCATCAGGCGATTCAAGGTTT 60
QY 322 GCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGCCACGTA 381
DB 61 ACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGCCACGTA 120
QY 382 AAGTAATGTTGAGGCTACTGGCTGCTACTATTGAGGCAAGGAGGATTTTCAAAATATG 441
DB 121 AAGAAGGTTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 442 GAGTAATCACCAAGGATAGAGACCCCAAGCAACCAAAATACAGCATGACTACTAC 501
DB 181 GAGTAATCACCAAGGATAGAGACCCCAAGCAACCAAAATACAGCATGACTACTAC 240
QY 502 AACAGTCCACAGAAAGTGGTACTGACAAATATTTCTAAGGGTGTTCCTGCTGATAGTGA 561
DB 241 AACAGTCCACAGAAAGTGGTACTGACAAATATTTCTAAGGGTGTTCCTGCTGATAGTGA 300

QY 435 AAATATGGAGATACACACACAGTGTAGAACCCCAAGCAACCAAAATACAGGCATGA 494
DB 1 AAATATGGAGATACACCGCAGTGGTAGAACCCCAAGCAACCAAAATACAGGCATGA 60
QY 495 CTACTACACAGTGCACAGAGTGGTACTGACAAATATTTCTAAGGGTGTTCCTGCTGA 554
DB 61 CTACTACACAGTGCACAGAGTGGTGTGACAAATATTTCTAAGGGTGTTCCTGCTGA 120
QY 555 TAGTGTAGTCAATTTTGTGTAACAGATGTTGAGTGTATCCATTTCAAGTTCCTGCTGA 614
DB 121 TAGTGTAGTCAATTTTGTGTAACAGATGTTAAGTGTATCCATTTCAAGTTCCTGCTGA 180
QY 615 AGAACCATACATTTTCAGCCCGCTGTTTCTTAAGATTTATCCCTGAGAAATGCAATA 674
DB 181 AGAACCATACATTTTCAGCCCGCTGTTTCTTAAGATTTATCCCTGAGAAATGCAATA 240
QY 675 TCAGTCTTATCCACCAAGGTTGAAATACGCTTGCACAAAGCTGACAGGTTGACATGAC 734
DB 241 TCAGTCTTATCCACCAAGGTTGAAATACGCTTGCACAAAGCTGACAGGTTGACATGAC 300
QY 735 AACCTGTGATTTATAGTGAAGACCAAGACTGTTCCCAAGAGATTAAGCAGCCAGCTGA 794
DB 301 AACCTGTGATTTATAGTGAAGACCAAGCTGTTCCCAAGAGATTAAGCAGCCAGCTGA 360
QY 795 AACAGCCCAAGACTTTCATCCCATCTTCAAGAGCGGAAAAAGACTGGGATAAATGGA 854
DB 361 AACAGCCCAAGACTTTCATCCCATCTTCAAGAGCGGAAAAAGACTGGGATAAATGGA 420
QY 855 AGCTGAAGTCAAAAGGAGGAGGAGGAGAAATCTGATGCTGATGCTGATGCAATGAACA 914
DB 421 AGCTGAAGTCAAAAGGAGGAGGAGGAGAAATCTGATGCTGATGCTGATGCAATGAACA 480
QY 915 ATTCTTCGTCATCTACAGAGTGTCTGATGAAGATATGCGGAGGCGCATGATGAAGTC 974
DB 481 ATTCTTCGTCATCTACAGAGTGTCTGATGAAGATATGCGGAGGCGCATGATGAAGTC 540
QY 975 ATTCTGGAATCAATGTCATGTTCTTCAACCAATTTGAAAGATGTTGAGCAAAAGAA 1034
DB 541 ATTCTGGAATCAATGTCATGTTCTTCAACCAATTTGAAAGATGTTGATCAAAAGGT 600
QY 1035 GGTAGAA-GGAGAGCCCTGATGGTATGAGCTCAAGAGTGGGAATACTAA 1086
DB 601 GGTAGAAAGGAGCCCTCTGATGGTATGGAAGCTTAANAAGTGGGAATACTAA 653

RESULT 12
CA229980

QY 562 GTCAATGATTTTGGTGAACAGAGTGTAGTGTATCCATTGAAGTTCCTGGTGAAGACCA 621
 |||||
 Db 301 GTCAATGATTTTGGTGAACAGAGTGTAGTGTATCCATTGAAGTTCCTGGTGAAGACCA 360
 |||||
 QY 622 TACCAATTTTCCGCGCTGTTTCTAAGATTATCCCTGAGAAATGCAATATCAAGTC 681
 |||||
 Db 361 TACCAATTTTCCGCGCTGTTTCTAAGATTATCCCTGAGAAATGCAATATCAAGTC 420
 |||||
 QY 682 TTATCCACCAAGGTTGAATACGCTTGCAAAAGCTGAGCAGGTGACATGACACACCTG 741
 |||||
 Db 421 TTATCCACCAAGTGAATACGCTTGCAAAAGCTGAGCAGGTGACATGACACACCTG 480
 |||||
 QY 742 GATTATAGTGAAGACCAAGACTGTTCCCAAGAGATAAGCACCCAGCTGAAACAGCC 801
 |||||
 Db 481 GATTATAGTGAAGACCAAGACTGTTCCCAAGAGATAAGCACCCAGCTGAAACAGCC 540
 |||||
 QY 802 CCAAGACCTTCATACCCATCTTCAAGGCGAAGAAAGACTGGATAAACTGGAGCTGAA 861
 |||||
 Db 541 CCAAGACCTTCATACCCATCTTCAAGGCGAAGAAAGACTGGATAAACTGGAGCTGAA 600
 |||||
 QY 862 GTCAAAAAGGAGGAGGAAG 883
 |||||
 Db 601 GTCAAAAAGGAGGAGGAAG 622
 |||||

RESULT 13
 AW042392
 LOCUS 61402B03.v1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
 DEFINITION mRNA sequence.
 ACCESSION AW042392
 VERSION AW042392.1 GI:5901292
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 AUTHORS Walbot, V.
 JOURNAL
 COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 614028 row: B column: 03.
 Location/Qualifiers

1. 586
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultur="W23"
 /db_xref="taxon:4577"
 /tissue_type="root"
 /dev_stage="3-4 days old"
 /lab_host="XL0LR"
 /clone_lib="614 - root cDNA library from Walbot Lab"
 /note="Organ: root; Vector: pBluescriptII SK+; Site: 1:
 EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
 Lab (LM)"
 ORIGIN
 Query Match 53.7%; Score 583.4; DB 2; Length 586;
 Best Local Similarity 99.7%; Pred. No. 2.6e-152;
 Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 147 GCTCGCAACTACACTGAGGCTGTGGCGGATGCTTACAAAGCAATTGAGCTTGATCTAT 206
 Db 1 GCTCGCAACTACACTGAGGCTGTGGCGGATGCTTACAAAGCAATTGAGCTTGATCTAT 60

QY 207 GATCATAAAGCTTACTACCGAAAGTGTGCTATGATTAAGCTTGAAGATACCAAC 266
 |||||
 Db 61 GATCATAAAGCTTACTACCGAAAGTGTGCTATGATTAAGCTTGAAGATACCAAC 120
 |||||
 QY 267 TGCAAGGCTGCTCTTGGAGTTGGTCTTCTTATGCTATCAGGCGATTCAAGGTTTGCTCG 326
 |||||
 Db 121 TGCAAGGCTGCTCTTGGAGTTGGTCTTCTTATGCTATCAGGCGATTCAAGGTTTGCTCG 180
 |||||
 QY 327 TCTATTGAAGATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGACCCAGTAAGAA 386
 |||||
 Db 181 TCTATTGAAGATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGACCCAGTAAGAA 240
 |||||
 QY 387 TGTGAGGCTACTGCTGCTACTATTGAGGCAAGGAGATTTCACAAATATGGAGAA 446
 |||||
 Db 241 TGTGAGGCTACTGCTGCTACTATTGAGGCAAGGAGATTTCACAAATATGGAGAA 300
 |||||
 QY 447 TACACCAAGGATGATGAACCCCAAGCAACCAAAATACAGGATGACTACTACACAG 506
 |||||
 Db 301 TACACCAAGGATGATGAACCCCAAGCAACCAAAATACAGGATGACTACTACACAG 360
 |||||
 QY 507 TGCCACAGAGTGTGCTACTGACATATTTGCTAAGGTTGTTCTGCTGATAGTAGTCAT 566
 |||||
 Db 361 TGCCACAGAGTGTGCTACTGACATATTTGCTAAGGTTGTTCTGCTGATAGTAGTCAT 420
 |||||
 QY 567 TGATTTTGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGACCATACCA 626
 |||||
 Db 421 TGATTTTGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGACCATACCA 480
 |||||
 QY 627 TTTTCAGCCCGTCTGTTTCTAAGATTATCCCTGAGAAATGCAAAATATCAAGTCTTATC 686
 |||||
 Db 481 TTTTCAGCCCGTCTGTTTCTAAGATTATCCCTGAGAAATGCAAAATATCAAGTCTTATC 540
 |||||
 QY 687 CACCAAGTGTGAATACGCTTCGAAAGCTGAGCAGGTGACATGG 732
 |||||
 Db 541 CACCAAGTGTGAATACGCTTCGAAAGCTGAGCAGGTGACATGG 586
 |||||

RESULT 14
 CA229238
 LOCUS

DEFINITION SCAGFL3023C08.g Saccharum officinarum FL3 Saccharum officinarum
 cDNA clone SCAGFL3023C08 5', mRNA sequence.
 CA229238
 CA229238.1 GI:35291554

EST.
 Saccharum officinarum
 Saccharum officinarum
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

REFERENCE 1. (bases 1 to 609)
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 The libraries that made SUCSEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccc.org.br
 Plate: 023 row: C column: 08
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers

FEATURES
 source

1. 609
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"

```

/clone="SCAGFL3023C08"
/lab host="DH10B"
/clone lib="Saccharum officinarum FL3"
/vector="Organ: Base of developing inflorescence (5cm-long) ;
Site 1: Sal1; Site 2: NotI; An
unidirectional cDNA library generated from [Base of
developing inflorescence (5cm-long)]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucst.lad.ic.uniscamp.br/public"

```

ORIGIN

Query Match	53.7%	Score 583.4	DB 6	Length 609
Best Local Similarity	97.4%	Pred. No. 2.6e-152		
Matches 593	Conservative 0	Mismatches 16	Indels 0	Gaps 0
Qy	262	CNAACTCGAAGGCTGCTCTTGAGTTCGGGTTCTTCCTATGCAATCAGGCGATTCAAGGTTT	321	
Db				
Qy	1	CNAACTCGAAGGCTGCTCTTGAGTTCGGGTTCTTCCTATGCAATCAGGCGATTCAAGGTTT	60	
Db				
Qy	322	GCTCGTCTATTGGAAGGAATGTGATGAGCGCATCGCTGAGGAACTCTAGCCAGGCCACCAAGTA	381	
Db				
Qy	61	ACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAACTCTAGCCAGGCCACCAAGTA	120	
Db				
Qy	382	AAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAAGGAGGATTTTCACAAAATATG	441	
Db				
Qy	121	AAGAAGGTTGAGGCTCTCTGTGGCTGCTACTGTTGAGGACAAAGGAGGATGCCACAAATATG	180	
Db				
Qy	442	GAGAAATACACCAACAGTGATAGAAACCCCAAGCAAAAACCAAAATACAGGCATGACTACTAC	501	
Db				
Qy	181	GAGAATACACCGCCAGTGCTGAGAAACCCCAAGCAAAAACCAAAATACAGGCATGACTACTAC	240	
Db				
Qy	502	AACAGTSCCAACAGAAAGTGCTACTGACAAATATTTGCTAAGGGGTGTTCTGCTGATAGTGTA	561	
Db				
Qy	241	AACAGTGCCAACAGAAAGTGCTGCTGACAAATATTTGCTAAGGGGTGTTCTGCTGATAGTGTA	300	
Db				
Qy	562	GTCAATTGATTTTGGTGAAACAGATGTTGAGTGATATCCATTGAAGTTCTTGGTGGAAGAACCA	621	
Db				
Qy	301	GTCAATTGATTTTGGTGAAACAGATGTTAAGTGATATCCATTGAAGTTCTTGGTGGAAGAACCA	360	
Db				
Qy	622	TACATTTTTCAGCCCGTCTGTTTTCTTAAGATATATCCCTGAGAAATGCAAAATATCAAATC	681	
Db				
Qy	361	TACATTTTTCAGCCCGTCTGTTTTCTTAAGATATATCCCTGAGAAATGCAAAATATCAAATC	420	
Db				
Qy	682	TTATCCACCAAGGTTGAAATACGGCTTGCACAAAGCTGAGCAGGTGCAATGCAACACCCCTG	741	
Db				
Qy	421	TTATCCACCAAGGTCGAATACGGCTTGCACAAAGCTGAGCAGGTGCAATGCAACACCCCTG	480	
Db				
Qy	742	GATTTATAGTGGAGACCAAGAGCTGTTTCCCAAGAGATTAAGCAGCGCTGAAACAGCC	801	
Db				
Qy	481	GATTTATAGTGGAGACCAAGAGCTGTTTCCCAAGAGATTAAGCAGCGCTGAAACAGCC	540	
Db				
Qy	802	CCAAGACCTTTCATACCAATCTTTCAAAGGCGCAAAAAAGACTGGGATATACTGGAAGCTGAA	861	
Db				
Qy	541	CCAAGACCTTTCATACCAATCTTTCAAAGGCGCAAAAAAGACTGGGATATACTGGAAGCTGAA	600	
Db				
Qy	862	GTCAAAAAG	870	
Db				
Qy	601	GTCAAAAAG	609	
Db				

RESULT 15

CA262395	CA262395	702 bp	linear	EST 26-SEP-2003
LOCUS	SCPRLEB2028A09.g	LB2 Saccharum officinarum	cdna clone	SCPRLEB2028A09
DEFINITION	5', mRNA sequence.			
ACCESSION	CA262395			
VERSION	CA262395.1	GI:35955513		
KEYWORDS	EST.			

```
Db 456 TCTAGCCAGGCCACCAAGTAAGAAGGTGAGGCTCCTGTGGTGCTACTGTTGAGGACAAG 515
QY 424 GAGGATTTCAAAATATGGAGATACACCACCAAGTATAGAACCCCAAGCAAAACCAAAA 483
Db 516 GAGGATGCCACAATATGGAGATACACGCCCAAGTGGTAGAACCCCAAGCAAAACCTAAA 575
QY 484 TACAGGCATGACTACTACAACAGTCCCAAGAGTGGTACTGACAATATTTGCTAAGGGT 543
Db 576 TACAGGCATGACTACTACAACAGTCCCAAGAGTGGTACTGACAATATTTGCTAAGGGT 635
QY 544 GTTCCTGCTGATAGTGTAGTCAATTCATTTGGTGAACAGATGTTGAGTGTATCCATTGAA 603
Db 636 GTTCCTGCTGATAGTGTANTCATTCATTTGGTGAACAGATGTTAAATGTATCCATTGAA 695
QY 604 GTTCCT 609
Db 696 GTNCCT 701
```

Search completed: April 23, 2005, 05:43:24
Job time : 3906.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 03:11:06 ; Search time 685.535 Seconds

(without alignments)

9617.545 Million cell updates/sec

Title: US-10-609-078-9

Perfect score: 1086

Sequence: 1 atggcgctgcggtctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084.4	99.9	1444	17	US-10-425-114-15440 Sequence 15440, A
2	1084.4	99.9	1450	17	US-10-425-114-25244 Sequence 25244, A
3	1084.4	99.9	1755	18	US-10-425-115-18040 Sequence 18040, A
4	988.2	91.9	1537	18	US-10-767-701-13951 Sequence 13951, A
5	977.6	90.0	1361	17	US-10-425-114-25845 Sequence 25845, A
6	977.6	90.0	1490	17	US-10-425-114-22598 Sequence 22598, A
7	977.6	90.0	1519	18	US-10-425-115-18041 Sequence 18041, A
8	862	79.4	1649	17	US-10-425-114-24871 Sequence 24871, A
9	862	79.4	1706	18	US-10-425-115-18039 Sequence 18039, A
c 10	736.6	67.8	2471	18	US-10-739-930-4858 Sequence 4858, Ap
11	698.2	64.3	1569	18	US-10-437-963-37229 Sequence 37229, A

12	490	45.1	571	16	US-10-074-473-2	Sequence 2, Appli
13	476.8	43.9	1691	17	US-10-424-599-65699	Sequence 55699, A
14	474.4	43.7	1598	17	US-10-424-599-81927	Sequence 81927, A
15	416.6	38.4	483	17	US-10-424-599-118531	Sequence 118531, A
16	332.8	30.6	1674	17	US-10-424-599-98272	Sequence 98272, A
17	286	26.3	287	9	US-09-294-0938-571	Sequence 571, App
18	243.4	22.4	1029	17	US-10-424-599-81933	Sequence 81933, A
19	221.8	20.4	585	18	US-10-021-323-2492	Sequence 2492, Ap
20	217.8	20.1	534	18	US-10-021-323-14417	Sequence 14417, A
21	216.4	19.9	272	9	US-09-294-0938-4704	Sequence 4704, Ap
22	201.6	18.6	851	17	US-10-424-599-75856	Sequence 75856, A
23	194.4	17.9	528	18	US-10-021-323-14351	Sequence 14351, A
c 24	183.6	16.9	573	18	US-10-021-323-2546	Sequence 2546, Ap
25	182.8	16.8	275	18	US-10-425-115-60448	Sequence 60448, A
26	180.6	16.6	656	17	US-10-424-599-98270	Sequence 98270, A
c 27	175.4	16.2	523	18	US-10-021-323-15961	Sequence 15961, A
28	156.4	14.4	1264	17	US-10-264-049-918	Sequence 918, App
29	156.4	14.4	1366	16	US-10-287-218-33	Sequence 33, Appl
30	156.4	14.4	1366	18	US-10-474-291-33	Sequence 99, Appl
31	156.4	14.4	1555	18	US-10-335-053-99	Sequence 22330, A
32	156.4	14.4	1776	18	US-10-357-930-22330	Sequence 28174, A
33	156.4	14.4	1776	18	US-10-357-930-28174	Sequence 45, Appl
34	156.4	14.4	2053	10	US-09-983-802-45	Sequence 45, Appl
35	156.4	14.4	2053	10	US-09-984-490-45	Sequence 82, Appl
36	156.4	14.4	2053	11	US-09-973-278-82	Sequence 3135, Ap
c 37	148.6	13.7	291	9	US-09-294-0938-3135	Sequence 621, App
38	141.4	13.0	409	9	US-09-770-423-621	Sequence 929, App
39	129.2	11.9	481	10	US-09-770-961-929	Sequence 3, Appli
40	125	11.5	388	16	US-10-074-473-3	Sequence 25942, A
c 41	106.8	9.8	479	17	US-10-424-599-25942	Sequence 49636, A
c 42	106	9.8	432	17	US-10-424-599-49636	Sequence 41142, A
43	94.4	8.7	317	17	US-10-424-599-41142	Sequence 1154, Ap
44	93.2	8.6	672	17	US-10-260-238-1154	Sequence 43909, A
45	93.2	8.6	672	18	US-10-437-963-43909	

ALIGNMENTS

RESULT 1

US-10-425-114-15440
; Sequence 15440, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaka, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 15440

; LENGTH: 1444

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3059-061-D10_FLI

US-10-425-114-15440

Query Match 99.9%; Score 1084.4; DB 17; Length 1444;

Best Local Similarity 99.9%; Pred. No. 5.3e-313;

Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCGCGCTCGGATCTGGAGAGCAAGCCAGAGGCGCTTCGTCGACGACTTCGAG 60

Db 83 ATGCGCGCTCGGATCTGGAGAGCAAGCCAGAGGCGCTTCGTCGACGACTTCGAG 142

Qy 61 CTGCCACCGAGCTCTACAGCCAGCCATCGACGCCGGCCCGCCGACCTCTAT 120

143 CTGGCCACCGAGCTCTACAGCAGGCCATCGAGCGGGCCGCCACCGCCGACCTCTAT 202
121 GCGACCGCGCCAGCGCAATCAAGCTCGCAATACACTGAGCTGTGGCGATGCT 180
203 GCGACCGCGCCAGCGCAATCAAGCTCGCAATACACTGAGCTGTGGCGATGCT 262
181 AACAAAGCAATTGAGCTTGTATCTATGATGATCAATAAGCTTACTACCGGAAGTGTGCA 240
263 AACAAAGCAATTGAGCTTGTATCTATGATGATCAATAAGCTTACTACCGGAAGTGTGCA 322
241 TGCAATTAAGCTTGAAGATACCAAACTGCAAGGCTCTTCTGAGTTGGTCTTCTTAT 300
323 TGCATTAAGCTTGAAGATACCAAACTGCAAGGCTCTTCTGAGTTGGTCTTCTTAT 362
301 GCATCAGCGCAATCAAGCTTGTCTGTATTAAGGAATGTGATGAGCGCATCGCTGAG 360
383 GCATCAGCGCAATCAAGCTTGTCTGTATTAAGGAATGTGATGAGCGCATCGCTGAG 442
361 GAATCTAGCCAGGACCAAGTAAGATGTTGAGCTACTGTGCTGTACTATTGAGGAC 420
443 GAATCTAGCCAGGACCAAGTAAGATGTTGAGCTACTGTGCTGTACTATTGAGGAC 502
421 AAGGAGGATTTCAAAATATGGGAATACACCACTGATAGAACCCCAAGCAACCA 480
503 AAGGAGGATTTCAAAATATGGGAATACACCACTGATAGAACCCCAAGCAACCA 562
481 AAATACAGGATGACTACTACAGCTGCAAGAGTGTGATGAGCAATATTGCTAAG 540
563 AAATACAGGATGACTACTACAGCTGCAAGAGTGTGATGAGCAATATTGCTAAG 622
541 GGTGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
623 GGTGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
601 GAGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
683 GAGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 742
661 GAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTGCAAAATACGCTTGCAGAGCTGAG 720
743 GAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTGCAAAATACGCTTGCAGAGCTGAG 802
721 CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
803 CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 862
781 AGCAGCGCAGCTGAAACAGCGCCCAAGACCTTATCCACCAAGGTGCAAAATACGCTTGCAGAGCTGAG 840
863 AGCAGCGCAGCTGAAACAGCGCCCAAGACCTTATCCACCAAGGTGCAAAATACGCTTGCAGAGCTGAG 922
841 TGGGATAAATGGAAGCTGAAAGTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
923 TGGGATAAATGGAAGCTGAAAGTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 982
901 GCTGCAATGAAACAAATTTCTCGTGACATCTACAGGATGCTGATGAGATATCGGAGG 960
983 GCTGCAATGAAACAAATTTCTCGTGACATCTACAGGATGCTGATGAGATATCGGAGG 1042
961 GCCATGATGAGTCAATTTCTCGTGAGATCAAAATGAGCACTGTTCTCAACCAATGGAAGAT 1020
1043 GCCATGATGAGTCAATTTCTCGTGAGATCAAAATGAGCACTGTTCTCAACCAATGGAAGAT 1102
1021 GTTGAGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1103 GTTGAGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1162
1081 TACTAA 1086
1163 TACTAA 1168

US-10-425-114-25244
; Sequence 25244, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25244
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3732-024-C2_FLI
US-10-425-114-25244

Query Match 99.9%; Score 1084.4; DB 17; Length 1450;
Best Local Similarity 99.9%; Pred. No. 5.3e-313;
Matches 1085; Conservative 1; Indels 0; Gaps 0;

QY 1 ATGGCCGCTCGGATCTGAGAGCAAGGCCAAGGAGGCTTCTGTCGACGACGACCTTCTGAG 60
DB 89 ATGGCCGCTCGGATCTGAGAGCAAGGCCAAGGAGGCTTCTGTCGACGACGACCTTCTGAG 148
QY 61 CTGGCCACCGAGCTCTACAGCCAGGCGCATTCGACGCGGGCCGCCACCGCCGACCTCTAT 120
DB 149 CTGGCCACCGAGCTCTACAGCCAGGCGCATTCGACGCGGGCCGCCACCGCCGACCTCTAT 208
QY 121 GCGACCGCGCCAGCGGCACATCAAGCTCGCCCACTACACTGAGGCTGTGGCGATGCT 180
DB 209 GCGACCGCGCCAGCGGCACATCAAGCTCGCCCACTACACTGAGGCTGTGGCGATGCT 268
QY 181 AACAAAGCAATTGAGCTTGTATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 269 AACAAAGCAATTGAGCTTGTATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 328
QY 241 TGCAATTAAGCTTGAAGATACCAAACTGCAAGGCTGCTCTGAGTTGGTCTTCTTAT 300
DB 329 TGCAATTAAGCTTGAAGATACCAAACTGCAAGGCTGCTCTGAGTTGGTCTTCTTAT 388
QY 301 GCATCAGCGCAATCAAGCTTGTCTGTATTAAGGAATGTGATGAGCGCATCGCTGAG 360
DB 389 GCATCAGCGCAATCAAGCTTGTCTGTATTAAGGAATGTGATGAGCGCATCGCTGAG 448
QY 361 GAATCTAGCCAGGACCAAGTAAGATGTTGAGCTACTGTGCTGTACTATTGAGGAC 420
DB 449 GAATCTAGCCAGGACCAAGTAAGATGTTGAGCTACTGTGCTGTACTATTGAGGAC 508
QY 421 AAGGAGGATTTCAAAATATGGAATACACCACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 509 AAGGAGGATTTCAAAATATGGAATACACCACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 568
QY 481 AAATACAGGATGACTACTACCAAGTCCACAGAGTGGTACTGACAAATATTGCTAAG 540
DB 569 AAATACAGGATGACTACTACCAAGTCCACAGAGTGGTACTGACAAATATTGCTAAG 628
QY 541 GGTGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 629 GGTGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
QY 601 GAGTTCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 689 GAGTTCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 748
QY 661 GAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTGCAAAATACGCTTGCAGAGCTGAG 720

```
Db 749 GAGAAATGCAATATCAAGTCTTATCCACCAAGGTGGAATACGCTTGCAGAAAGCTGAG 808
Qy 721 CAGGTGACATGGACAAACCTCGATATATAGTGGAGACCAAAAGACTGTTCCCGCAGAGATA 780
Db 809 CAGGTGACATGGACAAACCTCGATATATAGTGGAGACCAAAAGACTGTTCCCGCAGAGATA 868
Qy 781 AGCAGCGCAGCTGAAACAGCCCGCAGACCTTATACCCATCTTCAAGGCGGCAAAAAGAC 840
Db 869 AGCAGCGCAGCTGAAACAGCCCGCAGACCTTATACCCATCTTCAAGGCGGCAAAAAGAC 928
Qy 841 TGGGATAAATCGGAAGCTGAAAGTCAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 929 TGGGATAAATCGGAAGCTGAAAGTCAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
Qy 901 GCTGCAATGAACAAATCTTCCGTGACATCTACAAGGATGCTGATGAAGATATCGGAGG 960
Db 989 GCTGCAATGAACAAATCTTCCGTGACATCTACAAGGATGCTGATGAAGATATCGGAGG 1048
Qy 961 GCCATGATGAAGTCAATTCGTGGAATCAATGGCACTGTTCTCTCAACCAATTTGGAAGAT 1020
Db 1049 GCCATGATGAAGTCAATTCGTGGAATCAATGGCACTGTTCTCTCAACCAATTTGGAAGAT 1108
Qy 1021 GTTGGAGCAAAAGAGTAGAAGGAGGAGCCCGCTGATGGTATGGAGCTCAAGAGTGGGAA 1080
Db 1109 GTTGGAGCAAAAGAGTAGAAGGAGGAGCCCGCTGATGGTATGGAGCTCAAGAGTGGGAA 1168
Qy 1081 TACTAA 1086
Db 1169 TACTAA 1174
```

RESULT 3

```
US-10-425-115-18040
; Sequence 18040, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18040
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116455C.1
US-10-425-115-18040
```

```
Query Match 99.9%; Score 1084.4; DB 18; Length 1755;
Best Local Similarity 99.9%; Pred. No. 5.8e-313;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCCCGCTCGGATCTGGAGCAGCAGCGCAAGGAGGCTTCGTGCGAGCAGCTTCGAG 60
Db 252 ATGGCCCGCTCGGATCTGGAGCAGCAGCGCAAGGAGGCTTCGTGCGAGCAGCTTCGAG 311
Qy 61 CTGGCCACCGAGCTCTACAGCAGCGCCATCGACGCGCGGCGCGCCCGCCGCGACCTCTAT 120
Db 312 CTGGCCACCGAGCTCTACAGCAGCGCCATCGACGCGGCGCGCCCGCCGCGACCTCTAT 371
Qy 121 GCCGACCGCGCCCGCGCGCACATCAAGCTCGGCACTACATGAGGCTGTGGCGGAGTCT 180
Db 372 GCCGACCGCGCCCGCGCGCACATCAAGCTCGGCACTACATGAGGCTGTGGCGGAGTCT 431
Qy 181 AACAAAGCAATTGAGCTTGATCTTATGATGATGAAGCTTACTACCGGAGAGGTCGCA 240
```

```
Db 432 AACAAAGCAATTGAGCTTGATCTTATGATGATCAAAAGCTTACTACCGGAAAGTGTGCA 491
Qy 241 TGCATTAAGCTTGAAGAAATACCAAACTGCAAAAGGCTGCTCTTGAGTTGGGTTCTTCTAT 300
Db 492 TGCATTAAGCTTGAAGAAATACCAAACTGCAAAAGGCTGCTCTTGAGTTGGGTTCTTCTAT 551
Qy 301 GCATCAGCGGATTCAGGTTTGTCTCTATTTGAAGGAATGCTGATGAGCGCATCGCTGAG 360
Db 552 GCATCAGCGGATTCAGGTTTGTCTCTATTTGAAGGAATGCTGATGAGCGCATCGCTGAG 611
Qy 361 GAATCTAGCCAGGACCAAGTAAGAAATGTTGAGGCTACTGTGGCTGTCTATTTGAGGAC 420
Db 612 GAATCTAGCCAGGACCAAGTAAGAAATGTTGAGGCTACTGTGGCTGTCTATTTGAGGAC 671
Qy 421 AAGGAGGATTTCACAAAATAGGAGAAATACACACAGTATAGAACCCCAAGCAAAACA 480
Db 672 AAGGAGGATTTCACAAAATAGGAGAAATACACACAGTATAGAACCCCAAGCAAAACA 731
Qy 481 AAATACAGGATGACTACTACACAGTGCCACAGAGTGGTACTGACAAATATTTGCTAAG 540
Db 732 AAATACAGGATGACTACTACACAGTGCCACAGAGTGGTACTGACAAATATTTGCTAAG 791
Qy 541 GGTGTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 792 GGTGTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
Qy 601 GAAGTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 852 GAAGTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
Qy 661 GAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTGAAATACGCTTGCAGAAAGCTGAG 720
Db 912 GAGAAATGCAAAATATCAAGTCTTATCCACCAAGGTGAAATACGCTTGCAGAAAGCTGAG 971
Qy 721 CAGGTGACATGGACAAACCTCGATATATAGTGGAGACCAAAAGACTGTTCCCGCAGAGATA 780
Db 972 CAGGTGACATGGACAAACCTCGATATATAGTGGAGACCAAAAGACTGTTCCCGCAGAGATA 1031
Qy 781 AGCAGCGCAGCTGAAAAGCAGCCCAAGACCTTTCATCCATCTTCAAAAGCGGAAAGAGAC 840
Db 1032 AGCAGCGCAGCTGAAAAGCAGCCCAAGACCTTTCATCCATCTTCAAAAGCGGAAAGAGAC 1091
Qy 841 TGGGATAAATCGGAAGCTGGAAGTCAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 1092 TGGGATAAATCGGAAGCTGGAAGTCAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1151
Qy 901 GCTGATTTGAACAAATCTTCCGTCGATCTTACAGGATGCTGATGAAGATATGCGGAGG 960
Db 1152 GCTGATTTGAACAAATCTTCCGTCGATCTTACAGGATGCTGATGAAGATATGCGGAGG 1211
Qy 961 GCCATGATGAAGTCAATTCGTGGAATCAAAATGGCAGCTGTTCTCTCAACCAATTTGAAAGAT 1020
Db 1212 GCCATGATGAAGTCAATTCGTGGAATCAAAATGGCAGCTGTTCTCTCAACCAATTTGAAAGAT 1271
Qy 1021 GTTGGAGCAAAAGAGTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
Db 1272 GTTGGAGCAAAAGAGTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1331
Qy 1081 TACTAA 1086
Db 1332 TACTAA 1337
```

RESULT 4

```
US-10-767-701-13951
; Sequence 13951, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
```



```

; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 13951
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS2718_1
; US-10-767-701-13951

Query Match          91.9%;   Score 998.2;   DB 18;   Length 1537;
Best Local Similarity 95.3%;   Pred. No. 3.3e-287;
Matches 1043;   Conservative 0;   Mismatches 43;   Indels 9;   Gaps 1;

QY      1  ATGCCCGCTCGGATCTTGGAGAGCAAGGCCAAGGAGGCGCTTCTGTCGACGACGACTTCGAG 60
DB      94  ATGCCCGCTCGGATCTTGGAGAGCAAGGCCAAGGAGGCGCTTCTGTCGACGACGACTTCGAG 153
QY      61  CTGGCCACCGAGCTCTACAGCCAGGCCCATCGACGCCGCCGCGGCCGCCACCGCCGACCTCTAT 120
DB      154  CTGGCCCGCGAGCTCTACCCAGGCCCATCGACGCCGCCGCGGCCGCCACCGCCGACCTCTAC 213
QY      121  GCCGACCGCGCCCGCCAGGCGCACATCAAGCTCGGCAACTACATGAGGCTGTGGCGGATGCT 180
DB      214  GCCGACCGCGCCCGCCAGGCGCACATCAAGCTCGGCAACTACATGAGGCTGTGGCGGATGCT 273
QY      181  AACAAAGCAATTGAGCTTGATCTTATGATGATCAAAAGCTTACTACCGGAAAAGGTGCTGCA 240
DB      274  AACAAAGCAATTGAACTTTGATCTTACGATGATCAAAAGCTTACTACCGGAAAAGGTGCTGCA 333
QY      241  TGCATTAACTTGAAGAAATACAAACTCGAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
DB      334  TGCATTAACTTGAAGAAATACAAACTCGAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 393
QY      301  GCATCAGCGCATTCAGAGTTTGCCTCTATTGAAGGAATGTGATGAGCGCATCGCTCAG 360
DB      394  GCATCAGCGCATTCAGAGTTTACTCGCTCTATTGAAGGAATGTGAAGAGCGCATTCGCTCAG 453
QY      361  GAATCTAGCCAGGACACAGTAAAGAACTGTGAGGCTACTGTG-----GCTGCTACT 411
DB      454  GAATCTAGCCAGGCTCCAGCAAGAACTGTGAGCCTCTGTGGCTGCTGCTGCTACT 513
QY      412  ATTGAGGACAAGGAGGATTTCAAAATATGGAGAAATACACCAAGTAGTAAAGAACCCCA 471
DB      514  GTTGAGGACAAGGAGGATGTCGAAATATGGAGAAATACACCGCAGTGGTAGAACCCCA 573
QY      472  AGCAAAACAAATACAGCGCATGACTTACTACACAGTGCACAGAGAGTGGTACTGACAATA 531
DB      574  AGCAAAACCTAAATAGTCCCATGACTACTACAAAGTGCACAGAGAGTGGTATTGACAATA 633
QY      532  TTTGCTAAAGGGTGTTCCTGCTGATAGTGTAGTCAATTGATTTTGGTGAACAGATGTGAGT 591
DB      634  TTTGCTAAAGGGTGTTCCTGCTGATAGTGTAGTCAATTGATTTTGGTGAACAGATGTGAGT 693
QY      592  GTATCCATTGAAGTTCTCGGTGAGAGAACATACCACTTTACGCCCGCTGTGTTTCTTAAG 651
DB      694  GTGTCATTGAAGTCCCTGGTGAAGAACCATACCACTTTTCAGCCCCGCTGTGTTGCTAAG 753
QY      652  ATTATCCCTGAAATATGAAAATATCAAGTCTTTATCACCAAGGTTGAAATAGCCCTTGCA 711
DB      754  ATTATCCCTGAGAAATGCAAAATATCAAGTCTTTGTCACCAAGSTCGAAATACGCCTTGCA 813
QY      712  AAAGCTGACGAGGTGACATGGACAACCCCTGGAAATTATAGTGGAGAGCAAAAGACTGTTCC 771
DB      814  AAAGCTGACGAGGTGACATGGACAACCCCTGGAAATTATAGCGGAGAACCAAGGCTGTTCC 873
QY      772  CAGAAGATAAGCACGCCAGCTGAAACAGCGCCCAAGACCTTCATACCCATCTTCAAAGGCG 831
DB      874  CAGAAGATAAGCACGCCAGCTGAAACAGCGCCCAAGACCTTCATACCCATCTTCAAAGGCA 933
QY      832  AAAAAAAGACTGGGATAACTGGAAGCTGAAGTCAAAAAGAGGAGGAAGGAAGAAAACCTT 891

```

361	Qy	GAATCTAGCCAGGCAACGAGTAAGAATGTTGAGGCTA	CTGTGGCTGCTACTATTGAGGAC	420
459	Db	GAATCTAGCCAGGCAACGAGTAAGAATGTCGAGGCTCTGTAGCTGCTACTGTGAGGAC	518	
421	Qy	AAGGAGGATTTACAAAATATGGAGATACACACAGTGCATAGAACCCCCCAAGCAAAACCA	480	
519	Db	AAGGAGGATGTCGAAAATATGGATTAATACCGCCAGTGGTAGAACCCCAAGCAAAACCT	578	
481	Qy	AAATACAGGCATGACTACTACAAACAGTGCACAGAAAGTGGTACTGACAATATTTGCTAAG	540	
579	Db	AAATATAGGCATGACTACTACAAACAGTGCACAGAAAGTGGTCTCACATATATGCTAAG	638	
541	Qy	GGTGTTCTGCTGATAGTGTAGTCAATTGATTTTGTGGAACAGATGTTAGTGTATCCATT	600	
639	Db	GGTGTTCTGCTGATAGTGTAGTCAATTGATTTTGTGTGATCAGATGTTAGTGTATCCATT	698	
601	Qy	GAAGTTCTGCTGTGAAGAACCATACCAATTTGAGCCCGCTCTGTTTCTAAGATTATCCCT	660	
699	Db	GAAGTACCTGTGTGAAGAACCCGTAACCAATTTGAGCCCGCTCTGTTTCTAAGATTATCCCT	758	
661	Qy	GAGAAATCCAAATATCAAGTCTTTATCCACCAAGGTTGAAATACGCTTCCAAAAGCTGAG	720	
759	Db	GAGAAATCCAAAGTATCAAGTTTTATCCACCAAGGTCGAAATACGCTTCCAAAAGCTGAG	818	
721	Qy	CAGGTGACATGGACAACCCCTGGATTATAGTGAAGACCAAGAAGCTGTTCCCCAGAGAATA	780	
819	Db	CAGGTGACATGGACAACCCCTGGATTATAGTGAAGACCAAGAAGCTTATTTCCCCAGAGAATA	878	
781	Qy	AGCAAGCCAGGTGAACCAAGCCCCCAAGACTTCATACCCATCTTCAAAGCGCAAAAAGAC	840	
879	Db	AGCAAGCCAGGTGAACCAAGCCCCCAAGACTTCATACCCATCTTCAAAGTCAAAAAAGGAC	938	
841	Qy	TGGGATAAACTGGGAGCTGAGTCAAAAAGGAGGAGGAAGAAACCTTGATGGTGTAT	900	
939	Db	TGGGATAAACTGGGAGCTGAGTCAAAAAGGAGGAGGAAGAAACCTTGAGAGGTGAT	998	
901	Qy	GCTGCATTGAACAAATCTTCCGTGCACATCTACAAGGATGCTGATGAAGATATGCGGAGG	960	
999	Db	GCTGCATTGAACAAAGTTCTTCCGTGATATCTACAAGGATGCTGATGAAGATATGCGGAGG	1058	
961	Qy	GCCATGATGAAGTCAATTCGTGGAATCAAAATGGCACTGTTCTCTCAACCAATTTGGAAGAT	1020	
1059	Db	GCCATGGACAAGTCAATTCAGGGAATCTAATGGTACCGTTCTCTCAACCAATTTGGAAGAT	1118	
1021	Qy	GTTGGAGCAAAAGAGGTAGAGGGAGGCCCCCTGATGGTATGGAGCTCAAGAAAGTGGGAA	1080	
1119	Db	GTTGGATCAAGAAGCGTGGAGCGAGCCCTCTGATGGTATGGAGCTCAAGAAAGTGGGAA	1178	

RESULT 6

```

US-10-425-114-22598
; Sequence 22598, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22598
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LTB3591-005-E5_FLI
US-10-425-114-22598

```


Db	691	TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACCAAGGAGATTTCACAAATATGAGAA	750	Matches	906;	Conservative	0;	Mismatches	174;	Indels	45;	Gaps	2;
Qy	447	TACACCAACAGTATGAGAACCCCAAGCAAAACCAAAATACAGGATGACTACTACACAG	506										
Db	751	TACACCAACAGTATGAGAACCCCAAGCAAAACCAAAATACAGGATGACTACTACACAG	810										
Qy	507	TGCCACAGAGTGTACTGCAATATTTGCTTAAGGGTGTTCCTGCTGATAGTGTAGTCAT	566										
Db	811	TGCCACAGAGTGTACTGCAATATTTGCTTAAGGGTGTTCCTGCTGATAGTGTAGTCAT	870										
Qy	567	TGATTTTGGTGAAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA	626										
Db	871	TGATTTTGGTGAAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA	930										
Qy	627	TTTTTCCGCGCTGTTTCTTAAGATATCCCTGAGAAATGCAAAATATCAAGTCTTATC	686										
Db	931	TTTTTCCGCGCTGTTTCTTAAGATATCCCTGAGAAATGCAAAATATCAAGTCTTATC	990										
Qy	687	CACCAAGGTTGAAATACGCTTTCAGAAAGCTGACAGTGAATGCAACCCCTGGATTA	746										
Db	991	CACCAAGGTTGAAATACGCTTTCAGAAAGCTGACAGTGAATGCAACCCCTGGATTA	1050										
Qy	747	TAGTGGAGACCAAGACTGTTCCTCCAGAGATGAACAGCCAGCTGAAACAGCCCCAAG	806										
Db	1051	TAGTGGAGACCAAGACTGTTCCTCCAGAGATGAACAGCCAGCTGAAACAGCCCCAAG	1110										
Qy	807	ACCTTCATACCATCTTCAAGGGCGMAAAGACTGGGATAACTGGAAGCTGAAGTCAA	866										
Db	1111	ACCTTCATACCATCTTCAAGGGCGMAAAGACTGGGATAACTGGAAGCTGAAGTCAA	1170										
Qy	867	AAAGGAGGAGGAAGGAAAGAACTTGATGTGTGCTGCTGCAATGGAACAAATCTTCCGTGA	926										
Db	1171	AAAGGAGGAGGAAGGAAAGAACTTGATGTGTGCTGCTGCAATGGAACAAATCTTCCGTGA	1230										
Qy	927	CATCTACAGAGTGTCTGATGAAGATATCGGAGGGCCATGATGAAGTCAATTCGTGGATC	986										
Db	1231	CATCTACAGAGTGTCTGATGAAGATATCGGAGGGCCATGATGAAGTCAATTCGTGGATC	1290										
Qy	987	AAATGGCACTCTCTCTCAACCAATGGAAAGATGTGGAGCAAGAGGTGAGAGGGAG	1046										
Db	1291	AAATGGCACTCTCTCTCAACCAATGGAAAGATGTGGAGCAAGAGGTGAGAGGGAG	1350										
Qy	1047	CCCCCTCATGTGATGAGCTCAAGAGTGGGAATACTAA	1086										
Db	1351	CCCCCTCATGTGATGAGCTCAAGAGTGGGAATACTAA	1390										

RESULT 10
US-10-739-930-4858/c
; Sequence 4858, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 4858
; LENGTH: 2471
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2471)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER396_1
US-10-739-930-4858

Query Match 67.8%; Score 736.6; DB 18; Length 2471;
Best Local Similarity 80.5%; Pred. No. 7.3e-209;

Qy 1042 GGGAGCCCCCTGATGCTCAAGAGTGGGTAATAA 1086
Db |||||
1202 GGAAGCCCTCTGATGGAATGGAGCTCAAGAGTGGGATATAA 1158
|||

RESULT 11

US-10-437-963-37229
; Sequence 37229, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 37229
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40977C.1
US-10-437-963-37229

Query Match 64.3%; Score 698.2; DB 18; Length 1569;
Best Local Similarity 79.0%; Pred. No. 1.7e-197;
Matches 889; Conservative 0; Mismatches 188; Indels 48; Gaps 3;

Qy 4 GCCCGCTCGATCTGGAGAGCAAGCCCAAGAGGCTTCGTCGAGCAGCACTTCGAGCTG 63
Db |||||
94 GCCCGCTCGATCTGGAGAGCAAGCCCAAGAGGCTTCGTCGAGCAGCACTTCGAGCTC 153
|||
64 GCCACCGAGCTCTACAGCAGCGCATCGACCCGGCGCCGACCGCGACCTCTATGCC 123
|||
154 GCCCGGAGCTCTACAGCAGCGCATCGAGCGCAGCCCGCCGACCGCGAGCTCTACGCC 213
|||
124 GACCGCGCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
Db |||||
214 GACCGCGCCAGGCGCCATATCAAGCTAGGCAACTACACTGAGGCTGTAGCTGATGCTAAC 273
|||
184 AAAGCAATTGAGCTTGATCTATGATGATCAAAAGCTTACTACCGAAAGGTGCTGATGC 243
Qy |||||
274 AAGGCCATTGAATTTGACCCCATCAATGCACAAGGCTTATCTTCGTAAGGCGCTGCAATG 333
|||
244 ATTAAGCTTGAAGAATACCAACTGCAAGGCTGCTCTTGAGTTGGGTCTCTTATGCA 303
Db |||||
334 ATAGACTGGAGGAGTATCAAACTGCAAAAGAGCTCTTGAATTTGGGTACTCTGTCGCA 393
|||
304 TCAGCGGATCAAGGTTGCTGCTATTGTAAGGAATGATGAGCGCATGCTGAGGAA 363
Qy |||||
394 TCTGCTGACTCAAGTTTACTCGCTTAATGAAGAGTGTGATGAGCGCATTTGCTGAGGAG 453
|||
364 TCTAGCGGCAACGATTAAGAATGTTGAGG-----CTACTGTGGTGTCT 408
Db |||||
454 CTTAGTGAAGTCCCTGTTTGAAGAGGCTGAAGATGGAGCAGCTGCGCCCTCTGTGCTTCT 513
|||
409 ACTATTGAGGACAGGAGGATTTACAAATATGAGATACACCACAGTATAGAACCC 468
Db |||||
514 TTGTGTGAGGAAAGGATGCTGCAAAATGATGATATACACCACCAATGGTGA----- 569
|||
469 CCAAGCAAAACAAATACAGCGATGACTTACTACCAACAGTGGCACAAGAGTGTGCTGACA 528
Qy |||||
570 --AGTGAAGCCAAATACAGGACGACTTCTCAACAGTGTCTACAGAGTTGTATTGACA 627
Db |||||

Qy 529 ATATTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAATTTTGGTGAACAGATGTTG 588
Db |||||
628 ATTTTTCGAAGGGTGTTCCTGCTGAGATGTTGTTGTTGATTTTGGTGAACAAATGTTA 687
|||
589 AGTGTATCCATTTGAAGTTTCCTGCTGGAAGAACCATACATTTTCAGCCCGCTGTTTCT 648
Qy |||||
688 AGTGTGCTGATTTGAAGTTCCTGCTGGAGAGGCGGTACCATTTTCAGCCTGCTGTTTCT 747
|||
649 AAGATTATCCCTGAGAAATGCAAAATNTCAAGTCTTATCCCAAGAGTTGAATAGCCCTT 708
Db |||||
748 AAGATCATCCCTGAGAAAGCAGATACCAAGTGTCTATCCACGAAGGTTTGAATTAAGACTG 807
|||
709 GCAAAAGCTGAGCAGGCTGACATGACCAACCTCGATTATAGTGAAGACCAAGAGCTGTT 768
Qy |||||
808 GCTAAGCTGAACAGATTACATGACCTCACTTATGATATAAAACCAAGAGCTGTT 867
|||
769 CCCCAGAGATTAAGCAGCG-----CAGCTGAACAGGCC 801
Db |||||
868 CCACAAAGATAATCCCTCCAGTTTACTGATAGCCCTTCTTTTATCAGCTGAATCGGCC 927
|||
802 CCAAGACTTCATACCCTCTTCAAGCGCAAAAGAGCTGGGATAAATCGAAGCTGAA 861
Qy |||||
928 CAGAGGCCATCATATCTCTTCTCAAAATCCAAAGAAAGACTGGGATAAATCGAAGCTGAA 987
|||
862 GTCAAAAGAGGAGAGAGGAAAGAACTTCTGATGCTGCTGATGCAATTCGAAACAAATCTTTC 921
Db |||||
988 GTTAAAGAGGAGAGAGGAGAGGAGCTTGAAGCGGATGCTGATGCAAAATTTTTC 1047
|||
922 CGTGACATCTACAGAGGATGCTGATGAAGATATCGGAGGCGCATGATGAAGCTCATTCGTG 981
Qy |||||
1048 CGTGACATCTACAGTGTGATGAAGACATGCGAGGCAATGATGAATCTTTTGT 1107
|||
982 GAATCAAAATGCACTGTTCTCTCAACCAATTTGGAAGAGTGTGGAGCAAAAGAGTAGAA 1041
Db |||||
1108 GAATCTAACGGTACTTCTGTCGACCAATTTGGAAGAGTGTGGCTCGAAGAGGTAGAG 1167
|||
1042 GGGAGCCCCCTGATGCTGATGAGCTCAAGAGTGGGAATACTAA 1086
Qy |||||
1168 GGAAGCCCACTGATGGATGGAGCTTAAAGAAATGGGAGTACTAA 1212
Db |||||

RESULT 12

US-10-074-473-2
; Sequence 2, Application US/10074473
; Publication No. US20030167504A1
; GENERAL INFORMATION:
; APPLICANT: Crane, Edmund H. III
; TITLE OF INVENTION: Maize Earl Interactor Polynucleotides
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: 35718/241314
; CURRENT APPLICATION NUMBER: US/10/074.473
; CURRENT FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/268,157
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Zea mays
US-10-074-473-2

Query Match 45.1%; Score 490; DB 16; Length 571;
Best Local Similarity 100.0%; Pred. No. 1.7e-135;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCCGCTCGATCTGGAGAGCAAGCCCAAGAGGCTTCGTCGAGCAGACTTCGAG 60
Db |||||
82 ATGCCCGCTCGATCTGGAGAGCAAGCCCAAGAGGCTTCGTCGAGCAGACTTCGAG 141
|||
61 CTGCCCAACCGAGCTCTACAGCAGGCGCATCGAGCCGGCGCCGACCGCGACCTCTAT 120
Qy |||||
142 CTGCCCAACCGAGCTCTACAGCAGGCGCATCGAGCCGGCGCCGACCGCGACCTCTAT 201
Db |||||

; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_44998C.1
US-10-424-599-81927

Query Match 43.7%; Score 474.4; DB 17; Length 1598;
Best Local Similarity 66.2%; Pred. No. 1.4e-130; Indels 10; Gaps 2;
Matches 718; Conservative 0; Mismatches 356;

```
Qy 7 CGCTCGGATCTGGAGAGCAAGCCCAAGGAGGCTTTCGTGCGACGAGCTTCGAGCTGGGCC 66
Db 145 GCTTCGGATCTTGAGCTTAAGGCCAAGAGGCTTTCGAAGATGATPACTATGATCTCGCC 204
Qy 67 ACCGAGCTCTACAGCCAGCCATCGACGCCGGGCCCGCCAGCCGCTCTATGCGGAC 126
Db 205 TACGACCTCTTAACTCAGCGCATTTGGTCTCAGCCCCCAACAACGCGAGACTATATGCTGAC 264
Qy 127 CGCGCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGATGCTTAACAAA 186
Db 265 CGTGGCAAGTCAACATCAAGTCAACAACTCAGGAGGCTGTTCGTGATGCAAAACAG 324
Qy 187 GCAATTGAGCTTGATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 246
Db 325 GCGATTGATGATCTCTCACTCAAAAGCATATTTGCGAAGGTACCGCATGATC 384
Qy 247 AAGCTTGAAGATACCAAACTGCAAGGCTCTCTTTGAGTTGGGTTCTTTATGCAATCA 306
Db 385 AAGCTTGAAGATATCAGACTGCTAAGCAGCTCTAGAGATGGGTGCTTCATTTGCTCCT 444
Qy 307 GCGCATCAAGGTTGCTCGTCTATTGAAGATGATGATGATGATGATGATGATGATGATGAT 366
Db 445 GGAGATCTAAATTTACTGATTTGATCAAGACTGCGATGAAGTATGCGAGAGATCT 504
Qy 367 AGCCAGGACCAAGTAAAGATTTGAGGCTACTGTGGCTGCTACTATTGAGGACCAAGAG 426
Db 505 GGTGTCATACCATACAGAGAGAGACACACAGGGTGTCTGTACAAAAGCTGTTGAG 564
Qy 427 GATTTCAAAATATGGAGAATACACCAAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 486
Db 565 G-----CAGAGATGATCTTCCAGAGCCACTCAGTAAACCGTGTAAACCTTAATATAC 618
Qy 487 AGGCATGACTACTACAAAGTGCACAGAGTGTGATGATGATGATGATGATGATGATGATGAT 546
Db 619 AGGCATGATTTCTACCAAGAACTGATGAATGGTTATTTACCATATTTGCAAAAGGCAAT 678
Qy 547 CCTGCTGATGTAGTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 606
Db 679 CCAGAGACGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 738
Qy 607 CCTGGTGAAGAACCATACATTTTCAGCCCGCTGCTGTTTCTTAAGATTTATCCCTGAGAAA 666
Db 739 CTTTGAAGATGCTTATGTTCTTCAACCTCGCTTATTTGGAAGATCATACCTCCCAA 798
Qy 667 TGCAATATCAAGTCTTATCCCAAGGTTGAAATACGCTTGCACAAAGCTGAGCAGGTG 726
Db 799 TGCCTGATGAAGTTTGTCCCAAAATTTGAAATTTGCTTTGCAAAAGCAGATCATATC 858
Qy 727 ACATGGACAACCTGGATTTATAGTGAAGACCAAGACTGTTCCCGCAGAGATTAAGACG 786
Db 859 CAATGGACATCTTAGAATTTCAACAGGGTAGCAGTTGACAGAGGTTTAGTGTGTTG 918
Qy 787 CCAGCTGAACAGCCCAAGACCTTCAACCCATCTTCAAAAGGCGAAAAA-----AGACTG 842
Db 919 CCAGTTCAAGAGGTGAAGAACTTATCCATCTCAAAACCGAAGAAATACAGATTG 978
Qy 843 GGAATAACTGAAGTGAAGTCAAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 902
Db 979 GGAATAAGCTTGAAGTCAAGTTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
Qy 903 TGCATTGAACAAATTTCTCGTGACATCTACAGGATGCTGATGAGATATGCGGAGGC 962
```

```
Db 1039 TCGTTGAACAAATTTTTCGTGATATATATCAAGATGAGATGAGGACACAAGAAGAGC 1098
Qy 963 CATGATGAGTCAATTCGTGGATCAATGGGACACTGTTCTCTCAACCAATTTGGAAGATGT 1022
Db 1099 AATGAGCAATCAATTTGTGGAGTCTAATGGAACAGTACTGTCTACAAAATGGAAGAAGT 1158
Qy 1023 TGGAGCAAGAAAGGTAGAAAGGAGCCCTGATGGTATGAGCTCAAGAAGTGGGAATA 1082
Db 1159 GGGATCAATGAGGTACAGGAAGTCTCTCTGATGGCATGAGTTAAAGAAATGGTAATA 1218
Qy 1083 CTAA 1086
Db 1219 TTGA 1222
```

RESULT 15

US-10-424-599-118531
; Sequence 118531, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 118531
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78041C.1
US-10-424-599-118531

Query Match 38.4%; Score 416.6; DB 17; Length 483;
Best Local Similarity 91.9%; Pred. No. 1.4e-113;
Matches 440; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy 283 GAGTGGGTCTCTTATGATCAGGCGATTCAGGTTTGTCTGCTATTGAGGAATGT 342
Db 5 GCGTCCGGTCTCTTATGATCAGGCGATTCAGGTTTGTCTGCTATTGAGGAATGT 64
Qy 343 GATCAGCGCATCGCTGAGGAATCTAGCCAGCACAGTAAAGAAATGTTGAGGCTACTGTG 402
Db 65 GATGAGTGCAATGCTGAGGAATCTAGCCAGCACAGTAAAGAAATGTTGAGGCTACTGTG 124
Qy 403 GCTGCTACTATTGAGGCAAGGAGGATTTCAAAATATGGAGATATACACCCAGTGATA 462
Db 125 GCTGCTACTATTGAGGCAAGGAGGATTTGCAAAATATGGAGATATACACCCAGTGATA 184
Qy 463 GAACCCCAAGCAACCAAAATACAGGATGACTACTACAAAGTGCACAGAGAGTGATA 522
Db 185 GAACCCCAAGCAACCAAAATATACGATGACTACTACAAAGTGCACAGAGAGTGATA 244
Qy 523 CTGCAATATTTGCTTAAGGGTTCCTGCTGATAGTGTAGTCAATTTGTTGGTGAACAG 582
Db 245 CTCAATATATGCTTAAGGGTTCCTGCTGATAGTGTAGTCAATTTGTTGGTGAACAG 304
Qy 583 ATGTTGAGTGTATCCATTGAAAGTTCTCGGTGAAGAACCATACATTTTTCAGCCCGCTG 642
Db 305 ATGTTGAGTGTATCCATTGAAAGTTCTCGGTGAAGAACCGTACCATTTTTCAGCCCGCTG 364
Qy 643 TTTTCTAAGATTAATCCCTGAGAAATGCAAAATATCAAGTCTTATCCCAAGAGTTGAATA 702
Db 365 TTATCTAAGATTAATCCCTGAGAAATGCAAAATATCAAGTCTTATCCCAAGAGTTGAATA 424
Qy 703 CGCCTTCCAAAGCTGAGGAGTGCATGACACCTGGATTTAGTGGAGACCAA 761
Db 425 GGCCTTGCAAAAGCTGAGGAGTGCATGACACCTTTGATTTATAGCGGAGACCAA 483
```

Search completed: April 23, 2005, 09:26:00
Job time : 688.535 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:26:41 ; Search time 43 Seconds
(without alignments)
626.705 Million cell updates/sec

Title: US-10-609-078-8

Perfect score: 1861

Sequence: 1 MAASDLSEKAEAFVDDDFE.....GAKKVEGSPDGMELKKWEY 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:**

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap:**
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap:**
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap:**
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap:**
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap:**
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	464	24.9	196	3	US-09-227-357-393
2	464	24.9	228	3	US-09-227-357-401
3	323	17.4	181	4	US-09-270-767-44320
4	244	13.1	388	4	US-09-248-796A-18350
5	206	11.1	65	4	US-09-248-796A-16433
6	197.5	10.6	1427	4	US-09-551-974A-97
7	197.5	10.6	1427	4	US-09-565-501A-97
8	197.5	10.6	1427	4	US-09-639-206A-97
9	197.5	10.6	1427	4	US-09-874-923-97
10	197.5	10.6	1641	4	US-09-551-974A-96
11	197.5	10.6	1641	4	US-09-565-501A-96
12	197.5	10.6	1641	4	US-09-639-206A-96
13	197.5	10.6	1641	4	US-09-874-923-96
14	192	10.3	494	4	US-09-517-779-2
15	192	10.3	499	4	US-09-949-016-7370
16	181.5	9.8	563	4	US-09-949-016-7419
17	179.5	9.6	982	4	US-09-551-974A-95
18	179.5	9.6	982	4	US-09-565-501A-95
19	179.5	9.6	982	4	US-09-639-206A-95
20	179.5	9.6	982	4	US-09-874-923-95
21	178	9.6	146	3	US-09-227-357-396
22	174	9.3	332	4	US-09-949-016-7518
23	174	9.3	382	3	US-09-301-978C-2
24	174	9.3	546	3	US-08-533-669A-2
25	174	9.3	546	3	US-09-183-861-2
26	174	9.3	546	3	US-09-022-765-2
27	174	9.3	546	4	US-09-551-974A-2

28	174	9.3	546	4	US-09-565-501A-2
29	174	9.3	546	4	US-09-639-206A-2
30	174	9.3	546	4	US-09-874-923-2
31	174	9.3	546	4	US-08-798-841-2
32	164	8.8	144	3	US-09-045-973-9
33	163.5	8.8	346	4	US-09-538-092-723
34	158	8.5	508	4	US-09-639-207-1
35	153	8.2	494	4	US-09-270-767-43457
36	151	8.1	124	4	US-09-513-999C-8104
37	149.5	8.0	459	4	US-09-248-796A-14438
38	143.5	7.7	391	4	US-09-248-796A-16378
39	140	7.5	303	3	US-09-045-973-1
40	140	7.5	668	4	US-09-949-016-11278
41	139.5	7.5	484	2	US-08-879-260-4
42	139.5	7.5	484	3	US-09-231-529-4
43	139.5	7.5	484	3	US-08-977-816-4
44	139.5	7.5	484	4	US-09-639-207-5
45	137	7.4	369	4	US-09-538-092-1166

ALIGNMENTS

RESULT 1

US-09-227-357-393
; Sequence 393, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18

Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 9, Appli
Sequence 723, App
Sequence 1, Appli
Sequence 43457, A
Sequence 8104, Ap
Sequence 14438, A
Sequence 16378, A
Sequence 1, Appli
Sequence 11278, A
Sequence 4, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 1166, Ap

```

; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 393
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-227-357-393

Query Match      24.9%; Score 464; DB 3; Length 196;
Best Local Similarity 46.0%; Pred. No. 6.6e-37;
Matches 92; Conservative 41; Mismatches 61; Indels 6; Gaps 2;

QY 161 KYRHDYNSATEVLTITFAKGVDPADSVVIDFGQMLSVSIEVPGPEYHFOPLRFSKLIIP 220
DB 2 KIKDWYQTESQVITLMIKNVQKNDVNVFSEKLSALVKLPSCDYNKLELHPIIP 61
QY 221 EKCKTQVLSKTVEIRLAKAEQVTTTLDYSGRPKTPVQKISTPASTAPRPSYPPSKA-KK 279
DB 62 EQSTPKVLSTKIEIKKPEAVRWEKLEGQDGVPTPKQFVAD-----VKNLYPSSSPYTR 116
QY 280 DWDKLEAEVKEKEKEKELGDGAALNKFRRDIYKDADEDMRAMMKSFVESNGTVLSTNWK 339
DB 117 NWDKLVGTEIKBEKNEKLEKGDGAALNRLFOQIYSDGSDGVKRAMNKSFMESGGTVLSTNWS 176
QY 340 DYCAKKVEGSPDGMELKKW 359
DB 177 DVGKRVKEINPPDDMEWKY 196

RESULT 2
US-09-227-357-401
; Sequence 401, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 401
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-227-357-401

Query Match      24.9%; Score 464; DB 3; Length 228;
Best Local Similarity 46.0%; Pred. No. 8.3e-37;
Matches 92; Conservative 41; Mismatches 61; Indels 6; Gaps 2;

QY 161 KYRHDYNSATEVLTITFAKGVDPADSVVIDFGQMLSVSIEVPGPEYHFOPLRFSKLIIP 220
DB 34 KIKDWYQTESQVITLMIKNVQKNDVNVFSEKLSALVKLPSCDYNKLELHPIIP 93
```



```
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551,974A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-551-974A-97

Query Match      10.6%; Score 197.5; DB 4; Length 1427;
Best Local Similarity 27.3%; Pred. No. 9.1e-10;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

Qy      3 ASDLESKAKAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db      566 AKQKDEGNQYFKEDKPEAAVTEAIKRNPAEHTSYSNRAAYIKLGFANDALKDAEK 625

Qy      63 AIELDPMMHKAYYRKGACIKLEBYQTAKAALELG-----SSVASGDSRFRARLLKECDE 117
Db      626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYVTIMK----- 680

Qy      118 IAESSQAPVKNVEAT-----VAATIED-----KEDFTNMENPPVIEPPSPKPY 162
Db      681 IQEMASGQSADGDEAARRAMDDEPEIAAIMQDSYMLVLKE-----MONDPTRIQ-----EY 731

Qy      163 RHDYNSATEVLTIFAKGVADSVWIDFGQMLSVSEVPGEPYHFQPLFSKIPEK 222
Db      732 MKD--SGISSKINKLISAG-----IIRFGQFSLTDPAVLGEETH-----LRVRVDPK 778

Qy      223 CKYQVLSTKVEIRLAKAEQV 242
Db      779 ANKLTIVEDNGIGMTKADLV 798

RESULT 7
US-09-565-501A-97
; Sequence 97, Application US/0956501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565,501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-565-501A-97

Query Match      10.6%; Score 197.5; DB 4; Length 1427;
Best Local Similarity 27.3%; Pred. No. 9.1e-10;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

Qy      3 ASDLESKAKAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db      566 AKQKDEGNQYFKEDKPEAAVTEAIKRNPAEHTSYSNRAAYIKLGFANDALKDAEK 625

Qy      63 AIELDPMMHKAYYRKGACIKLEBYQTAKAALELG-----SSVASGDSRFRARLLKECDE 117
Db      626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYVTIMK----- 680

Qy      118 IAESSQAPVKNVEAT-----VAATIED-----KEDFTNMENPPVIEPPSPKPY 162
Db      681 IQEMASGQSADGDEAARRAMDDEPEIAAIMQDSYMLVLKE-----MONDPTRIQ-----EY 731

Qy      163 RHDYNSATEVLTIFAKGVADSVWIDFGQMLSVSEVPGEPYHFQPLFSKIPEK 222
Db      732 MKD--SGISSKINKLISAG-----IIRFGQFSLTDPAVLGEETH-----LRVRVDPK 778

Qy      223 CKYQVLSTKVEIRLAKAEQV 242
Db      779 ANKLTIVEDNGIGMTKADLV 798

RESULT 7
US-09-565-501A-97
; Sequence 97, Application US/0956501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565,501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-565-501A-97
```

```
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-565-501A-97

Query Match      10.6%; Score 197.5; DB 4; Length 1427;
Best Local Similarity 27.3%; Pred. No. 9.1e-10;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

Qy      3 ASDLESKAKAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db      566 AKQKDEGNQYFKEDKPEAAVTEAIKRNPAEHTSYSNRAAYIKLGFANDALKDAEK 625

Qy      63 AIELDPMMHKAYYRKGACIKLEBYQTAKAALELG-----SSVASGDSRFRARLLKECDE 117
Db      626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYVTIMK----- 680

Qy      118 IAESSQAPVKNVEAT-----VAATIED-----KEDFTNMENPPVIEPPSPKPY 162
Db      681 IQEMASGQSADGDEAARRAMDDEPEIAAIMQDSYMLVLKE-----MONDPTRIQ-----EY 731

Qy      163 RHDYNSATEVLTIFAKGVADSVWIDFGQMLSVSEVPGEPYHFQPLFSKIPEK 222
Db      732 MKD--SGISSKINKLISAG-----IIRFGQFSLTDPAVLGEETH-----LRVRVDPK 778

Qy      223 CKYQVLSTKVEIRLAKAEQV 242
Db      779 ANKLTIVEDNGIGMTKADLV 798

RESULT 8
US-09-639-206A-97
; Sequence 97, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-639-206A-97

Query Match      10.6%; Score 197.5; DB 4; Length 1427;
Best Local Similarity 27.3%; Pred. No. 9.1e-10;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

Qy      3 ASDLESKAKAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db      566 AKQKDEGNQYFKEDKPEAAVTEAIKRNPAEHTSYSNRAAYIKLGFANDALKDAEK 625

Qy      63 AIELDPMMHKAYYRKGACIKLEBYQTAKAALELG-----SSVASGDSRFRARLLKECDE 117
Db      626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYVTIMK----- 680

Qy      118 IAESSQAPVKNVEAT-----VAATIED-----KEDFTNMENPPVIEPPSPKPY 162
Db      681 IQEMASGQSADGDEAARRAMDDEPEIAAIMQDSYMLVLKE-----MONDPTRIQ-----EY 731
```

Qy 163 RHDYNSATEVLTIFAKGVADSVVDFGQMLSVSEVPGPEYHFPQPLFSKIPEK 222
Db 732 MKD--SGISSKINKLISAG-----IIRFGQFSLTDPVILGEETH-----LRVRVDPK 778
Qy 223 CKYQVLSTKVEIRLAKAEQV 242
Db 779 ANKLTVDNGIGMTKADLV 798

RESULT 9

US-09-874-923-97
; Sequence 97, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874.923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-874-923-97

Query Match 10.6%; Score 197.5; DB 4; Length 1427;
Best Local Similarity 27.3%; Pred. No. 9.1e-10;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

Qy 3 ASDLESKAKEAFVDDDFELATELYSQADGAPATADLYADRAQAHIKLGNYTEAVADANK 62
Db 566 AKQKDEGNQYFKEDKPEAAVATEAKRNPAAHTSYSNRAAYIKLGAFNDAKDAEK 625
Qy 63 AIELDPMMHKAYYKGAACIKLEBYOTAKALELG-----SSYASGDSRFARLLKECDE 117
Db 626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK- 680
Qy 118 IAESSQAPVKNVEAT-----VAATIED-----KEDFTNMENTPPVIEPPSKPKY 162
Db 681 IQEMASGQSDGDEAARRAMDPEIAIMQSYNQLVLKE-----MNDPTRIQ-----EY 731
Qy 163 RHDYNSATEVLTIFAKGVADSVVDFGQMLSVSEVPGPEYHFPQPLFSKIPEK 222
Db 732 MKD--SGISSKINKLISAG-----IIRFGQFSLTDPVILGEETH-----LRVRVDPK 778
Qy 223 CKYQVLSTKVEIRLAKAEQV 242
Db 779 ANKLTVDNGIGMTKADLV 798

RESULT 10

US-09-551-974A-96
; Sequence 96, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551.974A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-551-974A-96

Query Match 10.6%; Score 197.5; DB 4; Length 1641;
Best Local Similarity 27.3%; Pred. No. 1.1e-09;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;
Qy 3 ASDLESKAKEAFVDDDFELATELYSQADGAPATADLYADRAQAHIKLGNYTEAVADANK 62
Db 566 AKQKDEGNQYFKEDKPEAAVATEAKRNPAAHTSYSNRAAYIKLGAFNDAKDAEK 625
Qy 63 AIELDPMMHKAYYKGAACIKLEBYOTAKALELG-----SSYASGDSRFARLLKECDE 117
Db 626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK- 680
Qy 118 IAESSQAPVKNVEAT-----VAATIED-----KEDFTNMENTPPVIEPPSKPKY 162
Db 681 IQEMASGQSDGDEAARRAMDPEIAIMQSYNQLVLKE-----MNDPTRIQ-----EY 731
Qy 163 RHDYNSATEVLTIFAKGVADSVVDFGQMLSVSEVPGPEYHFPQPLFSKIPEK 222
Db 732 MKD--SGISSKINKLISAG-----IIRFGQFSLTDPVILGEETH-----LRVRVDPK 778
Qy 223 CKYQVLSTKVEIRLAKAEQV 242
Db 779 ANKLTVDNGIGMTKADLV 798

RESULT 11

US-09-565-501A-96
; Sequence 96, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565.501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-565-501A-96

Query Match 10.6%; Score 197.5; DB 4; Length 1641;

This Page Blank (uspio)
- 2 -

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:17:15 ; Search time 164 Seconds
(without alignments)
851.344 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESAKEAFVDDPE.....GAKKVBGSPDGMELKKWEY 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: Geneseq_1980s:*
- 2: Geneseq_1990s:*
- 3: Geneseq_2000s:*
- 4: Geneseq_2001s:*
- 5: Geneseq_2002s:*
- 6: Geneseq_2003s:*
- 7: Geneseq_2003bs:*
- 8: Geneseq_2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1589.5	85.4	367	8	ADQ37151 OSGT1 pr
2	1589.5	85.4	367	8	ADQ15735 Rice stre
3	1135	61.0	358	3	AG21460 Arabidops
4	1135	61.0	358	3	AGS2816 Arabidops
5	1134	60.9	350	3	AGS29759 Arabidops
6	1134	60.9	371	3	AGS29758 Arabidops
7	1128	60.6	358	3	AGS24735 Arabidops
8	870	46.7	278	3	AGS52817 Arabidops
9	870	46.7	278	3	AGS21461 Arabidops
10	870	46.7	278	3	AGS24736 Arabidops
11	857	46.1	270	3	AGS29760 Arabidops
12	791	42.5	250	3	AGS24737 Arabidops
13	791	42.5	250	3	AGS21462 Arabidops
14	791	42.5	250	3	AGS52818 Arabidops
15	615	33.0	333	6	ABO07188 Human p53
16	615	33.0	333	6	ADJ66517 Suppresso
17	615	33.0	344	5	ABP41961 Human ova
18	594	31.9	365	5	ABP09748 Amino aci
19	594	31.9	365	6	ABP58341 Human cel
20	508.5	27.3	110	7	ABM74423 DNA clone
21	489	26.3	252	3	AAI73366 HTRM clon
22	464	24.9	196	2	AAI02827 Fragment
23	464	24.9	196	7	ADA07572 Human sec
24	464	24.9	196	8	ADN41381 Novel hum
25	464	24.9	228	7	ADA07580 Human sec

26	464	24.9	228	8	ADN41389	Adn41389 Novel hum
27	445	23.9	413	5	ABP73229	Abp73229 Candida a
28	405	21.8	400	6	ABJ25759	Abj25759 Aspergill
29	383	20.6	478	6	ABJ26359	Abj26359 Aspergill
30	317	17.0	178	4	ABB64536	Abb64536 Drosophil
31	311.5	16.7	283	4	AAM90770	Aam90770 Human imm
32	300	16.1	368	7	ADF75113	Adf75113 A_gosaypi
33	224	12.0	482	5	AAE16244	Aae16244 Tomato in
34	224	12.0	485	5	AAE16236	Aae16236 Tomato in
35	203.5	10.9	955	7	ADB78890	Adb78890 Leishmani
36	200.5	10.8	369	3	AAG05677	Agag05677 Arabidops
37	199.5	10.7	714	4	ABG09988	Abg09988 Novel hum
38	197.5	10.6	1437	5	AAU71857	Aau71857 Leishmani
39	197.5	10.6	1427	5	AAU71311	Aau71311 MAPS1A-M1
40	197.5	10.6	1427	7	ADB78860	Adb78860 Leishmani
41	197.5	10.6	1641	5	AAU71856	Aau71856 Leishmani
42	197.5	10.6	1641	5	AAU71310	Aau71310 MAPS1A-M1
43	197.5	10.6	1641	7	ADB78859	Adb78859 Leishmani
44	192	10.3	494	5	ABG70122	Abg70122 Human pre
45	192	10.3	499	4	AAE09732	Aae09732 Protein p

ALIGNMENTS

RESULT 1
ADQ37151
ID ADQ37151 standard; protein; 367 AA.

XX ADQ37151;

XX AC

XX 07-OCT-2004 (first entry)

XX OSGT1 protein.

XX cell proliferation related polypeptide; cell proliferation; senescence;

XX differentiation; stress response.

XX Oryza sativa.

XX WO2004061122-A2.

XX 22-JUL-2004.

XX 23-DEC-2003; 2003WO-US041200.

XX 26-DEC-2002; 2002US-0436565P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Cooper B;

XX WPI; 2004-534388/51.

XX New nucleic acid molecule encoding a cell proliferation-related polypeptide, useful for modulating cell proliferation, senescence, differentiation, development, and stress response in plants, and for producing enhanced food crops.

XX Claim 1; SEQ ID NO 310; 408pp; English.

XX The present invention relates to an isolated nucleic acid molecule encoding a cell proliferation-related polypeptide. The nucleic acid molecule and the encoded polypeptide, and methods are useful for modulating cell proliferation, senescence, differentiation, development, and stress response in plants, and for producing enhanced food crops. The present sequence represents a cell proliferation-related polypeptide. The present sequence is published separately from the main body of the specification as EPO data.

XX Sequence 367 AA;

Query Match 85.4%; Score 1589.5; DB 8; Length 367;

```
Best Local Similarity 83.8%; Pred. No. 1.7e-137;
Matches 306; Conservative 28; Mismatches 24; Indels 7; Gaps 2;

QY 2 AASDLESKAEAFVDDDFELATLYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
DB 5 AASDLESKAAAFVDDDFELAAELYTQAEASPATAEALYADRAQAHIKLGNYTEAVADAN 64
QY 62 KAEIDPMMHKAYYRKGAAICIKLEEVYOTAKAALELGSSVYASGDSRPARLLKCEDERIAEE 121
DB 65 KAEIDPMSHKAYLRKGAACIRLEEVYOTAKAALELGSPASGDSRFRLLMKCEDERIAEE 124
QY 122 SSOAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKYRHDYNSATEVLT 176
DB 125 LTFVPPYKKAEDGAAASVASFVEEKDDAANMDNTPPMVE--VKPKYRHDYNSATEVLT 182
QY 177 IFAGKVPADSVVIDFGEOMLSVSIIEVPGEPYHFQRLFSKIIPKCKYQVLSTKVEIRL 236
DB 183 IFAGKVPANVVVDVFEQMLSVSIEVPGEPYHFQRLFSKIIPKSKYQVLSTKVEIRL 242
QY 237 AKAEQVTTWTLDSYGRPKTVPOKISTPAETAPRPSYPSSKAKKDWKLEAEVKKEEKEK 296
DB 243 AKAEQITWTLSDYDKKPKAVPOKIIPPAESAQRPSYPSSKSKKDWKLEAEVKKEEKEK 302
QY 297 LDGDAALNKFFRDIYKDAEDMRRAMKSFVESNGTVLSTNWKDVGAKKVEGSPDGMEL 356
DB 303 LEGDAALNKFFRDIYSDADEMRRAMKSFVESNGTVLSTNWKDVGSKKVEGSPDGMEL 362
QY 357 KKWEY 361
DB 363 KKWEY 367

RESULT 2
ADQ15735
ID ADQ15735 standard; protein; 367 AA.
XX
AC ADQ15735;
XX
DT 07-OCT-2004 (first entry)
XX
DE Rice stress-related protein #72.
XX
KW rice; stress-related protein; plant maturation; plant development;
KW plant proliferation; plant senescence; plant disease-resistance;
KW plant stress response; transgenic plant; pest tolerance;
KW herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
KW improved nutritional value; increased yield; increased proliferation.
XX
OS Oryza sativa.
XX
PN WO2004061080-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041098.
XX
PR 26-DEC-2002; 2002US-0436564P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
XX WPI; 2004-534374/51.
DR N-PSDB; ADQ15734.
XX
XX New isolated nucleic acids and proteins, useful for producing transgenic
PT plants having improved properties, e.g. tolerance to pests, herbicides,
PT or biotic or abiotic stresses, improved nutritional value, or increased
PT yield or proliferation.
XX
XX Claim 27; SEQ ID NO 144; 551pp; English.
PS
XX The invention comprises the amino acid and coding sequences of rice
CC
```

```
CC stress-related proteins. The DNA and protein sequences of the invention
CC are useful for regulating and controlling plant maturation and
CC development, including proliferation, senescence, disease-resistance, or
CC stress response. They are also useful for producing transgenic plants
CC having improved properties, e.g. tolerance to pests, herbicides, or
CC biotic or abiotic stresses, improved nutritional value, increased yield
CC or proliferation, or improved structure causing less loss from lodging or
CC shattering. The present amino acid sequence represents a rice stress-
CC related protein of the invention.
XX
SQ Sequence 367 AA;
Query Match 85.4%; Score 1589.5; DB 8; Length 367;
Best Local Similarity 83.8%; Pred. No. 1.7e-137;
Matches 306; Conservative 28; Mismatches 24; Indels 7; Gaps 2;

QY 2 AASDLESKAEAFVDDDFELATLYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
DB 5 AASDLESKAAAFVDDDFELAAELYTQAEASPATAEALYADRAQAHIKLGNYTEAVADAN 64
QY 62 KAEIDPMMHKAYYRKGAAICIKLEEVYOTAKAALELGSSVYASGDSRPARLLKCEDERIAEE 121
DB 65 KAEIDPMSHKAYLRKGAACIRLEEVYOTAKAALELGSPASGDSRFRLLMKCEDERIAEE 124
QY 122 SSOAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKYRHDYNSATEVLT 176
DB 125 LTFVPPYKKAEDGAAASVASFVEEKDDAANMDNTPPMVE--VKPKYRHDYNSATEVLT 182
QY 177 IFAGKVPADSVVIDFGEOMLSVSIIEVPGEPYHFQRLFSKIIPKCKYQVLSTKVEIRL 236
DB 183 IFAGKVPANVVVDVFEQMLSVSIEVPGEPYHFQRLFSKIIPKSKYQVLSTKVEIRL 242
QY 237 AKAEQVTTWTLDSYGRPKTVPOKISTPAETAPRPSYPSSKAKKDWKLEAEVKKEEKEK 296
DB 243 AKAEQITWTLSDYDKKPKAVPOKIIPPAESAQRPSYPSSKSKKDWKLEAEVKKEEKEK 302
QY 297 LDGDAALNKFFRDIYKDAEDMRRAMKSFVESNGTVLSTNWKDVGAKKVEGSPDGMEL 356
DB 303 LEGDAALNKFFRDIYSDADEMRRAMKSFVESNGTVLSTNWKDVGSKKVEGSPDGMEL 362
QY 357 KKWEY 361
DB 363 KKWEY 367

RESULT 3
AAG21460
ID AAG21460 standard; protein; 358 AA.
XX
AC AAG21460;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24024.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
```

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132248P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01334256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140951P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.

PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	23-OCT-1999;	99US-0161404P.	
PR	23-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161320P.	
PR	28-OCT-1999;	99US-0161392P.	
PR	28-OCT-1999;	99US-0161393P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 61.0%; Score 1135; DB 3; Length 358;			
Best Local Similarity 61.0%; Pred. No. 1.3e-95;			
Matches 224; Conservative 49; Mismatches 76; Indels 18; Gaps 4;			
QY	3	ASDLESKAEAFVDDDFELATELYSOAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62	
Db	2	AKELAEKAEAFVDDDFVADVLYSKAIDLDPNCAFFADRAQANIKINDFTEAVVDANK 61	
QY	63	AIELDPMHKAHYRKGAACIKLEEYQAKAALBELSSYASGDSRFARLLKCEDERIAEBS 122	
Db	62	AIELEPTLAKAYLRKGTACMKLESEYSTAKAALKEGASVAPNEPKFKMIDCDLRAEEB 121	
QY	123	SQ-----APVKVEATVAATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYNSATEVV 174	
Db	122	KDVLQPMPPSPSSSTPLATEAD-----APFVPIPAAPAKPMFRHEFYQKPEBAV 172	
QY	175	LTIFAKGVPSADSVVIDFGEOMLSVSTEVPGEPPYHFQPRLFSKIIEPKKYOVLSKVEI 234	
Db	173	VTIFAKVPKENVTFEFGQILSVVIDVAGEEAYHLOPRLFGKLIIEKRCRFVLSKVEI 232	
QY	235	RLAKAEQVTTTLDYSGRPKTVPKISTPAETAPEFSYSSKAKDMDKLEAEVKKKEEKE 294	
Db	233	RLAKAEIITWASLEY-GKGSVLPKPNVSSALSQRFPYSSKPAKDWKLEAEVKKQKXD 291	
QY	295	EKLDGDAALNKFPRDIYKDAEDMRAMMKS FVESNGTVLSTNWKDVGAKKVGSPDGM 354	
Db	292	EKLDGDAAMNKFSDIYSSADEDMRAMMKSFAESNGTVLSTNWKVEGVTKKVESTPPDGM 351	
QY	355	ELKKWEY 361	
Db	352	ELKKWEY 358	
RESULT 4			
AAG52816			
ID	AAG52816 standard; protein; 358 AA.		
XX	AC AAG52816;		
XX	DT 18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 67180.		
XX	Protein identification; signal transduction pathway; metabolic pathway;		

KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	Arabidopsis thaliana.
OS	EP1033405-A2.
XX	06-SEP-2000.
PN	25-FEB-2000; 2000EP-00301439.
XX	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.
PR	18-JUN-1999; 99US-0139459P.
PR	18-JUN-1999; 99US-0139460P.
PR	18-JUN-1999; 99US-0139461P.
PR	18-JUN-1999; 99US-0139462P.
PR	18-JUN-1999; 99US-0139463P.
PR	18-JUN-1999; 99US-0139750P.
PR	21-JUN-1999; 99US-0139763P.
PR	21-JUN-1999; 99US-0139817P.
PR	22-JUN-1999; 99US-0139899P.
PR	23-JUN-1999; 99US-0140353P.
PR	23-JUN-1999; 99US-0140354P.

QY 355 ELKKWEY 361
Db 352 ELKKWEY 358

RESULT 5

AAG29759

ID : AAG29759 standard; protein; 350 AA.

XX AC

XX AAG29759;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 35462.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EF1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 15-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 18-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136332P.

PR 28-MAY-1999; 99US-0136782P.

PR 03-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

```
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      60.9%; Score 1134; DB 3; Length 371;
Best Local Similarity 61.4%; Pred. No. 1.7e-95;
Matches 221; Conservative 56; Mismatches 69; Indels 14; Gaps 4;

QY 3 ASDLESKAEAFVDDDFELATLYSOAIDAGPATADLYADRAQAHTKLGNYTEAVADANK 62
DB 23 AKELDADKAEAFVDDDFDVAVDYKAIIDLPNCABEFFADRAQAQYTKLESFTEAVADANK 82

QY 63 AIELDPMHKAYRYKGAACIKLEEYQTAKAALFELSSYASGDSRFARLLKCECDERIAERS 122
DB 83 AIELDPSLTAYLRYKCTACWGLEEYRTAKALEKGASITPSESKFKKLLIDECNFLTIEE 142

QY 123 SQAPVKVQVATVAATIEDKEDFTNMMENTPPVIEPPSKP--KYRHVYNSATEVVLTFIFAK 180
DB 143 KDL-VQVPSTLPSSV-----TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAK 191

QY 181 GVPADSVVDFEGOMLSVIEVPGEPYHFQPLFSKIIPEKCKYOVLSKVEIRLAKAE 240
DB 192 GIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPLFKGIIIPDKCKYEVLSKIEICLAKAD 251

QY 241 QVTWTLDYSGRPKTVQKISTPAETAPRPSYPSSKAKKDWKLEAEVKKKEEKLDCG 300
DB 252 IITWASLEHGKGPVLP-KPNVSSEVSQRPAPSSKKVQDWKLEAEVKKQEKLEGD 310

QY 301 AALNKFPRDIYKDADEMRMMKSPVESNGTVLSTNWKDVGAKKVGGSPDGMELKKWE 360
DB 311 AALNKFPRDIYQNADEMRMMKSPVESNGTVLSTNQEVGKTKTISTPPDGMELKKWE 370

RESULT 7
AAG24735
ID AAG24735 standard; protein; 358 AA.
XX AC AAG24735;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28523.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 23-MAR-1999; 99US-0123548P.
XX PR 25-MAR-1999; 99US-0125788P.
XX PR 29-MAR-1999; 99US-0126264P.
XX PR 01-APR-1999; 99US-0126785P.
XX PR 06-APR-1999; 99US-0127462P.
XX PR 08-APR-1999; 99US-0128234P.
XX PR 16-APR-1999; 99US-0128714P.
XX PR 19-APR-1999; 99US-0129845P.
XX PR 21-APR-1999; 99US-0130077P.
XX PR 23-APR-1999; 99US-0130449P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.

99US-0132486P.
99US-0132487P.
99US-0132863P.
99US-0134256P.
99US-0134218P.
99US-0134219P.
99US-0134221P.
99US-0134370P.
99US-0134768P.
99US-0134941P.
99US-0135124P.
99US-0135353P.
99US-0135629P.
99US-0136021P.
99US-0136392P.
99US-0137222P.
99US-0137282P.
99US-0137528P.
99US-0137502P.
99US-0137724P.
99US-0138094P.
99US-0138540P.
99US-0138847P.
99US-0139119P.
99US-0139452P.
99US-0139453P.
99US-0139492P.
99US-0139454P.
99US-0139455P.
99US-0139456P.
99US-0139457P.
99US-0139458P.
99US-0139459P.
99US-0139460P.
99US-0139461P.
99US-0139462P.
99US-0139463P.
99US-0139750P.
99US-0139763P.
99US-0139817P.
99US-0139899P.
99US-0140353P.
99US-0140354P.
99US-0140695P.
99US-0140823P.
99US-0140991P.
99US-0141287P.
99US-0141842P.
99US-0142154P.
99US-0142055P.
99US-0142390P.
99US-0142803P.
99US-0142920P.
99US-0142977P.
99US-0143542P.
99US-0143624P.
99US-0144005P.
99US-0144085P.
99US-0144086P.
99US-0144325P.
99US-0144331P.
99US-0144332P.
99US-0144333P.
99US-0144334P.
99US-0144335P.
99US-0144352P.
99US-0144632P.
99US-0144884P.
99US-0144814P.
99US-0145086P.
99US-0145088P.
99US-0145089P.
99US-0145087P.
```


PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0133256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152263P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

PR	06-OCT-1999;	99US-0157865P.	XX	EP1033405-A2.
PR	07-OCT-1999;	99US-0158029P.	PN	
PR	08-OCT-1999;	99US-0158232P.	XX	
PR	12-OCT-1999;	99US-0158369P.	PD	06-SEP-2000.
PR	13-OCT-1999;	99US-0159293P.	XX	
PR	13-OCT-1999;	99US-0159294P.	PF	25-FEB-2000; 2000EP-00301439.
PR	13-OCT-1999;	99US-0159295P.	XX	
PR	14-OCT-1999;	99US-0159329P.	PR	25-FEB-1999;
PR	14-OCT-1999;	99US-0159330P.	PR	99US-0121825P.
PR	14-OCT-1999;	99US-0159331P.	PR	99US-0123180P.
PR	14-OCT-1999;	99US-0159337P.	PR	99US-0123548P.
PR	14-OCT-1999;	99US-0159388P.	PR	99US-0125788P.
PR	18-OCT-1999;	99US-0159584P.	PR	99US-0126264P.
PR	21-OCT-1999;	99US-0160741P.	PR	99US-0126785P.
PR	21-OCT-1999;	99US-0160767P.	PR	99US-0128234P.
PR	21-OCT-1999;	99US-0160770P.	PR	99US-0128714P.
PR	21-OCT-1999;	99US-0160814P.	PR	99US-0128845P.
PR	21-OCT-1999;	99US-0160815P.	PR	99US-0130077P.
PR	21-OCT-1999;	99US-0160880P.	PR	99US-0130449P.
PR	22-OCT-1999;	99US-0160981P.	PR	99US-0130510P.
PR	22-OCT-1999;	99US-0160989P.	PR	99US-0130891P.
PR	25-OCT-1999;	99US-0161404P.	PR	99US-0131449P.
PR	25-OCT-1999;	99US-0161405P.	PR	99US-0132048P.
PR	25-OCT-1999;	99US-0161406P.	PR	99US-0132484P.
PR	26-OCT-1999;	99US-0161359P.	PR	99US-0132485P.
PR	26-OCT-1999;	99US-0161360P.	PR	99US-0132486P.
PR	26-OCT-1999;	99US-0161361P.	PR	99US-0132487P.
PR	28-OCT-1999;	99US-0161920P.	PR	99US-0132863P.
PR	28-OCT-1999;	99US-0161922P.	PR	99US-0134256P.
PR	28-OCT-1999;	99US-0161933P.	PR	99US-0134218P.
PR	29-OCT-1999;	99US-0162142P.	PR	99US-0134219P.
Query Match				99US-0134221P.
Best Local Similarity				99US-0134370P.
Matches 173; Conservative 37; Mismatches 60; Indels 18; Gaps 4;				99US-0134768P.
				99US-0134941P.
QY	82	IKLEEVOTAKALELSSVAGSDSRFARLLKCEDERIAEBSQ-----APVKQNVATVA 135	PR	99US-0135124P.
Db	1	MKLEEVSTAKALEKASVAPNEPKFKMLDECDLRAIEEKDLVQMPSPSLPSSSTTPL 60	PR	99US-0135233P.
QY	136	ATIEDKEDFTNMMENTPPVIEP--PSKPKYRHYNSATEVVLTIFAKGVPAQSVVDQGE 193	PR	99US-0135235P.
Db	61	ATEAD-----APPVPIPAAPAKPMERHEFYQKPEEAVVTIFAKVPKENVTVEFGE 111	PR	99US-0136021P.
QY	194	OMLSVIEVPGEPYHFORLFSKIIPECKYOVLTSTKVEIRLAKAEQVTTTLDYSGRP 253	PR	99US-0136392P.
Db	112	QILSVVIDVAGEBAYHLOPLFKIIPKCRFEVLSTKVEIRLAKAEIITWASLEY-GKG 170	PR	99US-0136782P.
QY	254	KTPVQKISTPAETAPRPSYPSSKAKDWDKLEAEVKKKEEKKLDGDAALNKFFRDIYKD 313	PR	99US-0137224P.
Db	171	QSVLPKPNVSSALSQRPVYSSPKPAQWDKLEAEVKKQEKDEKLDGDAAMKFFSDIYSS 230	PR	99US-0137528P.
QY	314	ADEMRAMKSVESNGTVLSTNWKDVGAKKVEGSPDGMELKKWEY 361	PR	99US-0137724P.
Db	231	ADEMRAMKSPAESNGTVLSTNWKVEVGTKKVESTPPDGMELKKWEY 278	PR	99US-0138094P.
RESULT 9				99US-0138540P.
AAG21461				99US-0138847P.
XX	AC	AAG21461, standard; protein; 278 AA.	PR	99US-0139119P.
XX	AC	AAG21461,	PR	99US-0139452P.
XX	DT	17-OCT-2000 (first entry)	PR	99US-0139453P.
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24025.	PR	99US-0139454P.
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;	PR	99US-0139455P.
KW	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	99US-0139456P.
KW	KW	termination sequence.	PR	99US-0139457P.
XX	OS	Arabidopsis thaliana.	PR	99US-0139458P.
			PR	99US-0139459P.
			PR	99US-0139460P.
			PR	99US-0139461P.
			PR	99US-0139462P.
			PR	99US-0139463P.
			PR	99US-0139750P.
			PR	99US-0139763P.
			PR	99US-0139817P.
			PR	99US-0139899P.
			PR	99US-0140353P.
			PR	99US-0140354P.
			PR	99US-0140695P.
			PR	99US-0140823P.
			PR	99US-0140991P.
			PR	99US-0141287P.

```
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 02-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
```

```
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 14-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match

46.7%; Score 870; DB 3; Length 278;

Best Local Similarity 60.1%;

Pred. No. 2.4e-71;

Matches 173; Conservative 37; Mismatches 60; Indels 18; Gaps 4;

```
QY 82 IKLEEYOTAKALELGSSVAGDSRPARLLKCEDERIAESSQ-----APVKNVETVA 135
Db 1 MKLEEYSTAKALEKAGSAPNPKPKMIDECDLRIAEKKDLVQMPPLSSSTTPL 60
QY 136 ATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYNSATEVLTIFAKGVPADSVWIDFGE 193
Db 61 ATEAD-----APPVPIPAAPAKPMFRHEFYQKPEEAVVTIFAKKVPKENVTVEFGE 111
QY 194 QMLSVSIEVPGSEPHFQPLRSKIIPEKCKYQVLSTKVEIRLAKAEQVTTWLDYSGRP 253
Db 112 QILSVVIDVAGBEAVHLQPLFKIIPKCRFEVLSTKVEIRLAKAEIITWASLEY-GKG 170
QY 254 KTVPOKISTPARTAPRPSYSSKAKKDWKLEAEVKEKEKKLDGDAALNKPFRTDYKD 313
Db 171 QSVLPKPNVSSALSQRVPVYPSKPKAKDWKLEAEVKKQEKDEKLDGDAANKKFFSDIYS 230
QY 314 ADEDMRRAMKSFVESNGTGLSTNWKDVGAKKVEGSPDPGMELKKWEY 361
Db 231 ADEDMRRAMKSFVESNGTGLSTNWKVEGTKKVESTPPDGMELKKWEY 278
```

RESULT 10

AAG24736

ID AAG24736 standard; protein; 278 AA.

XX

AC AAG24736;

XX


```
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151103P.
PR 31-AUG-1999; 99US-0151130P.
PR 01-SEP-1999; 99US-0151438P.
PR 07-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153070P.
PR 15-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.7%; Score 870; DB 3; Length 278;
Best Local Similarity 60.1%; Pred. No. 2.4e-71;
Matches 173; Conservative 37; Mismatches 60; Indels 18; Gaps 4;

QY 82 IKLEEYQTAALLEGSSVAGSDSRPARLLKCEDERIAESQ-----APVKNVETVA 135
DB 1 MKLEEYSTAKALEGASVAPNEPKFKMIDCDLRAIEEEKDLVQMPPLSFSSSTPL 60
QY 136 ATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYNSATEVLTIFAKGVPAQSVVIDFGE 193
DB 61 ATEAD-----APPVPIPAAPKPMFRHFEYQKPEEAVVTIFAKVFPKENVTVEFGE 111
QY 194 QMLSVSIEVGPBEPYHQRFLRSKIPEKCKYQVLSTKVEIRLAKAEQVTTTLTLDYSGRP 253
DB 112 QILSVVIDVAGEAYHLQRLFGKIPEKCRFEVLSTKVEIRLAKAEIITWASLEY-GKG 170
QY 254 KTVPOKISTPAETAPRPSVPSSKAKDWKLEAEVKKEBKELGDAALNKFPRDIYKD 313
DB 171 QSVLPKPNVSSALSORPVYFSSKPAKDWKLEAEVKKEBKELGDAALNKFPSDIYS 230
QY 314 ADEDMRRAMKSFVESNGTGLSTNNKVGAKKVEGSPDPGMELKKWEY 361
```

```
Db 231 ADEDMRRAMKSFABESNGTGLSTNNKVEGTKKVESTPPDGMELKKWEY 278
RESULT 11
AAG29760
ID AAG29760 standard; protein; 270 AA.
XX AC AAG29760;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 35463.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 16-APR-1999; 99US-0128714P.
XX PR 19-APR-1999; 99US-0129845P.
XX PR 21-APR-1999; 99US-0130077P.
XX PR 23-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 18-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137222P.
XX PR 03-JUN-1999; 99US-0137528P.
XX PR 04-JUN-1999; 99US-0137502P.
XX PR 07-JUN-1999; 99US-0137724P.
XX PR 08-JUN-1999; 99US-0138094P.
XX PR 10-JUN-1999; 99US-0138540P.
XX PR 10-JUN-1999; 99US-0138647P.
XX PR 14-JUN-1999; 99US-0139119P.
XX PR 16-JUN-1999; 99US-0139452P.
XX PR 16-JUN-1999; 99US-0139453P.
XX PR 17-JUN-1999; 99US-0139492P.
XX PR 18-JUN-1999; 99US-0139454P.
XX PR 18-JUN-1999; 99US-0139455P.
```

```
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149375P.

PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150568P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 857; DB 3; Length 270;
Best Local Similarity 59.4%; Pred. No. 3.6e-70;
Matches 167; Conservative 45; Mismatches 55; Indels 14; Gaps 4;

QY 82 IKLEEVOTAKAALGSSVAGSDSRFARLLKCEDRIAEESQAPVKNVEATVAATIEDK 141
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MKLEERYTAKTALKEGASITPSESFKKLIDECNFIIEEKDL-VQVPSTLPSSV--- 56
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 142 EDFTNNMTPPVIEPPSKP--KYRHDYNSATEVVLTIIPAKVDPADSVVIDGEQMLSVS 199
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 57 -----TAPPVSELDVTPAKYREHYOKPEEVVVVTFKGIKQKNNVIDGEQILSVV 109
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 200 IEVPGEEPYHFQRLFSKIIPEKCKYQVLTSTKVEIRLAKAEQVTTWTLDYSGRPKTPQK 259
```

Db 110 I E V P G E D A Y L Q P R L F G K I I P D K C K Y E V L S T K I E I C L A K A D I I T W A S L E H G K G P A V L P - K 168
Qy 260 I S T P A E T A P R P S Y P S K A K D W K L E A E V K K E E K L D G D A A L N K F R D I Y K D A D E D M R 319
Dy 169 P N V S E V S Q R P A P S K K V K D W K L E A E V K K E K D K L E G D A A L N K F R E I Y Q N A D E D M R 228
Qy 320 R A M K S F V E S N G T V L S T N K D V G A K K V E G S P D G M E L K K W E 360
Dy 229 R A M K S F V E S N G T V L S T N Q E V G T K T I E S T P D G M E L K K W E 269

RESULT 12

AAG24737

ID AAG24737 standard; protein; 250 AA.

AC AAG24737;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.

PR	09-AUG-1999;	99US-0147493P.	QY	110	LLKCEDERIAESSQ-----APVKNEATVAATIEDKEDFTNMENTPPVIEP--PSKPK	161
PR	09-AUG-1999;	99US-0147935P.			:::	
PR	10-AUG-1999;	99US-0148171P.	Db	1	MIDECOLRIAEKEDLVQMPESLFSSTTPLATEAD-----APPVPIPAAPAKPM	51
PR	11-AUG-1999;	99US-0148319P.			:::	
PR	12-AUG-1999;	99US-0148341P.	QY	162	YHDIYNSATEVLTIFAKGVPADSVVIDFGHOMLSVIEVGEPEPYHQPRLFSKIIP	221
PR	13-AUG-1999;	99US-0148565P.			:::	
PR	13-AUG-1999;	99US-0148684P.	Db	52	FRHEFYQPEEAVVTIFAKVPEKENVTFEGEILSVVIDVAGEAYHQLQPLFGKIIP	111
PR	16-AUG-1999;	99US-0149368P.			:::	
PR	17-AUG-1999;	99US-0149175P.	QY	222	KCKYOVLSKVEIRLAKAEQVTTTLDYSGRKTVPQKISTPAETAPRPSYSSKAKDM	281
PR	18-AUG-1999;	99US-0149426P.			:::	
PR	20-AUG-1999;	99US-0149722P.	Db	112	KCRFEVLSTKVEIRLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRVPVYSSKPAKM	170
PR	20-AUG-1999;	99US-0149723P.			:::	
PR	20-AUG-1999;	99US-0149929P.	QY	282	DKLEAEVKKKEEKLGDGAALNKFFRDIYKDADEDMRRAMMKSVESNGTVLSTNWKDV	341
PR	23-AUG-1999;	99US-0149902P.			:::	
PR	23-AUG-1999;	99US-0149930P.	Db	171	DKLEAEVKKQEKDEKLGDGAAMNKEFSDIYSSADEDMRRAMMKSVESNGTVLSTNWKDV	230
PR	25-AUG-1999;	99US-0150566P.			:::	
PR	26-AUG-1999;	99US-0150884P.	QY	342	GAKVGEPPDGMELKKWEY	361
PR	27-AUG-1999;	99US-0151065P.			:::	
PR	27-AUG-1999;	99US-0151066P.	Db	231	GTKVVESTPPDGMELKKWEY	250
PR	27-AUG-1999;	99US-0151080P.			:::	
PR	30-AUG-1999;	99US-0151303P.			:::	
PR	31-AUG-1999;	99US-0151438P.			:::	
PR	01-SEP-1999;	99US-0151930P.			:::	
PR	07-SEP-1999;	99US-0152363P.			:::	
PR	10-SEP-1999;	99US-0153070P.			:::	
PR	13-SEP-1999;	99US-0153758P.	AC		AAG21462;	
PR	15-SEP-1999;	99US-0154018P.	XX			
PR	16-SEP-1999;	99US-0154039P.	DT	17-OCT-2000	(first entry)	
PR	20-SEP-1999;	99US-0154779P.	XX			
PR	22-SEP-1999;	99US-0155139P.			:::	
PR	23-SEP-1999;	99US-0155486P.	XX			
PR	24-SEP-1999;	99US-0155659P.			:::	
PR	28-SEP-1999;	99US-0156458P.			:::	
PR	29-SEP-1999;	99US-0156596P.			:::	
PR	04-OCT-1999;	99US-0157117P.			:::	
PR	05-OCT-1999;	99US-0157753P.	OS		Arabidopsis thaliana.	
PR	06-OCT-1999;	99US-0157865P.	XX			
PR	07-OCT-1999;	99US-0158029P.	PN		EP1033405-A2.	
PR	08-OCT-1999;	99US-0158232P.	XX			
PR	12-OCT-1999;	99US-0158369P.			:::	
PR	13-OCT-1999;	99US-0159293P.	PD	06-SEP-2000.		
PR	13-OCT-1999;	99US-0159294P.	PF			
PR	13-OCT-1999;	99US-0159295P.	XX	25-FEB-2000;	2000EP-00301439.	
PR	14-OCT-1999;	99US-0159329P.	XX			
PR	14-OCT-1999;	99US-0159330P.	PR	25-FEB-1999;	99US-0121825P.	
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAR-1999;	99US-0123180P.	
PR	14-OCT-1999;	99US-0159337P.	PR	09-MAR-1999;	99US-0123548P.	
PR	14-OCT-1999;	99US-0159637P.	PR	23-MAR-1999;	99US-0125788P.	
PR	14-OCT-1999;	99US-0159638P.	PR	25-MAR-1999;	99US-0126264P.	
PR	18-OCT-1999;	99US-0159584P.	PR	29-MAR-1999;	99US-0126785P.	
PR	21-OCT-1999;	99US-0160741P.	PR	01-APR-1999;	99US-0127462P.	
PR	21-OCT-1999;	99US-0160767P.	PR	06-APR-1999;	99US-0128234P.	
PR	21-OCT-1999;	99US-0160768P.	PR	08-APR-1999;	99US-0128714P.	
PR	21-OCT-1999;	99US-0160770P.	PR	16-APR-1999;	99US-0129845P.	
PR	21-OCT-1999;	99US-0160814P.	PR	19-APR-1999;	99US-0130077P.	
PR	21-OCT-1999;	99US-0160815P.	PR	21-APR-1999;	99US-0130449P.	
PR	22-OCT-1999;	99US-0160980P.	PR	23-APR-1999;	99US-0130510P.	
PR	22-OCT-1999;	99US-0160981P.	PR	28-APR-1999;	99US-0130891P.	
PR	22-OCT-1999;	99US-0160989P.	PR	30-APR-1999;	99US-0131443P.	
PR	25-OCT-1999;	99US-0161404P.	PR	30-APR-1999;	99US-0132048P.	
PR	25-OCT-1999;	99US-0161405P.	PR	30-APR-1999;	99US-0132407P.	
PR	25-OCT-1999;	99US-0161406P.	PR	04-MAY-1999;	99US-0132484P.	
PR	26-OCT-1999;	99US-0161359P.	PR	05-MAY-1999;	99US-0132485P.	
PR	26-OCT-1999;	99US-0161360P.	PR	06-MAY-1999;	99US-0132486P.	
PR	26-OCT-1999;	99US-0161361P.	PR	06-MAY-1999;	99US-0132487P.	
PR	28-OCT-1999;	99US-0161920P.	PR	07-MAY-1999;	99US-0132863P.	
PR	28-OCT-1999;	99US-0161992P.	PR	11-MAY-1999;	99US-0134256P.	
PR	28-OCT-1999;	99US-0161993P.	PR	14-MAY-1999;	99US-0134218P.	
PR	29-OCT-1999;	99US-0162142P.	PR	14-MAY-1999;	99US-0134219P.	
PR			PR	14-MAY-1999;	99US-0134221P.	
PR			PR	18-MAY-1999;	99US-0134370P.	
PR			PR	19-MAY-1999;	99US-0134768P.	
PR			PR	20-MAY-1999;	99US-0134941P.	
PR			PR	20-MAY-1999;	99US-0135124P.	

Query Match 42.5%; Score 791; DB 3; Length 250;
Best Local Similarity 60.4%; Pred. No. 3.9e-64;
Matches 157; Conservative 32; Mismatches 53; Indels 18; Gaps 4;

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145228P.
PR 26-JUL-1999; 99US-0145278P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.5%; Score 791; DB 3; Length 250;
Best Local Similarity 60.4%; Pred. No. 3.9e-64;
Matches 157; Conservative 32; Mismatches 53; Indels 18; Gaps 4;

Qy 110 LLKCDERIAESSQ-----APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPK 161
Db 1 MIDECDLRIABEEDLVQMPSPSFPSSSTTPLATEAD-----APPVPIPAAPAKPM 51
Qy 162 YRHDYNSATEVLTITFAKGVADSVVIDFGQMLSVSIEVPGEPYHFQRLFSKIIE 221
Db 52 FRHEFYQKPEAVVTITFAKVPKENVTVEFGEILSVVIDVAGEAYHLQRLFGKIIE 111
Qy 222 KCKYQVLSTKVEIRLAKAEQVTTWTLDSGRPKTVPQKISTPAETAPRPSYPSSKAKDW 281
Db 112 KCRFEVLSTKVEIRLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRVPVYPSKPAKW 170
Qy 282 DKLEAEVKKEKEKLDGDAALNKFFRDIYKDADEDMERAMKSFVESNGTVLSTNWKDV 341
Db 171 DKLEAEVKKEKLDGDAALNKFFSDIYSSADEDMRRAMKSPAESNGTVLSTNWKDV 230
Qy 342 GAKKVEGSPDGMELKKWEY 361
Db 231 GTKKVESTFPDGMELKKWEY 250

RESULT 15
ABO07188
ID ABO07188 standard; protein; 333 AA.
XX ABO07188;
XX DT 13-AUG-2003 (first entry)
XX DE Human p53 modifying protein, SEQ ID 148.
XX KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
XX KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
XX KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
XX KW apoptotic disorder; cell proliferation disorder.
XX OS Homo sapiens.
XX FN WO2002099122-A1.
XX PD 12-DEC-2002.
XX PF 03-JUN-2002; 2002WO-US017382.
XX PR 05-JUN-2001; 2001US-0296076P.
XX PR 10-OCT-2001; 2001US-0328605P.
XX PR 15-FEB-2002; 2002US-0357253P.
XX PA (EXEL-) EXELIXIS INC.
XX FI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX DR WPI; 2003-156859/15.
XX DR N-PSDB; ACD13363.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:18:25 ; Search time 40 Seconds
(without alignments)
868.356 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLSKAKEAFVDDDFE.....GAKKVEGSPDGMELKKWEY 361
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1135	61.0	358	2	T13017
2	1134	60.9	350	2	T05589
3	383	20.6	198	2	T20305
4	364.5	19.6	395	2	S66940
5	341	18.3	444	2	T40307
6	327	17.6	469	2	T48783
7	223	12.0	533	2	E84858
8	193.5	10.4	572	2	E86257
9	192	10.3	498	1	S52570
10	187	10.0	358	2	T00954
11	186	10.0	499	1	A55346
12	181.5	9.8	543	2	A38093
13	180.5	9.7	473	1	T40391
14	177.5	9.5	589	2	A32567
15	176	9.5	569	2	S56658
16	174	9.3	349	2	T08782
17	174	9.3	526	1	T45058
18	169.5	9.1	513	1	S52571
19	163.5	8.8	346	2	S61991
20	163.5	8.8	591	2	T41531
21	161.5	8.7	558	2	T48150
22	160	8.6	627	2	T04562
23	159.5	8.6	479	1	T46576
24	159.5	8.6	591	2	T51996
25	157.5	8.5	901	2	JC7111
26	151	8.1	677	2	T45682
27	148.5	8.0	358	2	T37805
28	147	7.9	317	2	T37851
29	146.5	7.9	385	2	S40699

30 145.5 7.8 422 2 T24865
31 141.5 7.6 1979 2 JH0059
32 139 7.5 605 2 AF1917
33 139 7.5 781 2 F86457
34 137.5 7.4 547 2 AE1884
35 136.5 7.3 222 2 F86424
36 136 7.3 225 2 AE2539
37 132.5 7.1 320 2 T03899
38 130.5 7.0 297 2 T34141
39 130.5 7.0 1621 2 A82255
40 130 7.0 619 2 A36682
41 130 7.0 2025 2 JC5020
42 130 7.0 2055 2 T31110
43 129.5 7.0 427 2 T00960
44 129.5 7.0 542 2 AD1333
45 127.5 6.9 961 2 T32493

ALIGNMENTS

RESULT 1

T13017
hypothetical protein F8L21.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13017
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mei
submitted to the Protein Sequence Database, July 1999
A:Reference number: 217587
A:Accession: T13017
A:Molecule type: DNA
A:Residues: 1-358 <BEV>
A:Cross-references: UNIPROT:Q9SUT5; EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.50
C:Experimental source: cultivar Columbia; BAC clone F8L21
C:Genetics:
A:Gene: ATSP:F8L21.50
A:Map position: 4
A:Introns: 54/1; 77/2; 119/1; 161/2; 193/3; 215/3; 262/1; 288/3; 323/3

Query Match 61.0%; Score 1135; DB 2; Length 358;
Best Local Similarity 61.0%; Pred. No. 1.5e-69;

Matches 224; Conservative 49; Mismatches 76; Indels 18; Gaps 4;

Qy 3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKGNVTEAVADANK 62
Db 2 AKELAERAKAEFLDDDFDVAVDLYSKAIDLDPNCAAFPADRAQANIKIDNFEAVVDANK 61
Qy 63 AIELDPMHKAHYKKGACIKLEEVQAKALELGSVSGDSRPARLLKCEDRIASES 122
Db 62 AIELEPTLAKAYLRKGTACMKLEYSYAKAALKEGASVAPNEPKFGMIDCDLRIABEE 121
Qy 123 SQ-----APVKNEATVAATIEDKDFTNMENTPPVIEP--PSKPKYRHDYVNSATEVV 174
Db 122 KDLVQPMPPSLPSSSTTPLATEAD-----APPVIPAAPAKPMFEHFYQKPEAV 172
Qy 175 LTIFAKGVADSVVDIFGEQMLSVSIEVPGEPVHQPRLFSKIIPKCKYQVLSKVEI 234
Db 173 VTIFAKKVPKENVTVEFGEQILSVVIDVAGEAYHLQPRLFKTIPEKCRPEVLSKVEI 232
Qy 235 RLAKAEQVTTLDYSGRPKTVPOKISTPATAPRPSYPSSKAKKDWDLAEVKKEKE 294
Db 233 RLAKAEIITWASLY-KGGSVLPKPNVSSALSQRVPVYPSKPAKDWDLAEVKKQSKD 291
Qy 295 EKLGDGAALNKKFFRDIYKDADEDMRRAMNKS FVSNGTVLSTNNKDVGAKEVSGSPDGM 354
Db 292 EKLGDGAAMNKKFFSDIYSSADEDMRRAMNKS FALSNGTIVLSTNNKEVGTKKVSTPPDGM 351
Qy 355 ELKKWEY 361
Db 352 ELKKWEY 358

```
RESULT 2
T05589
hypothetical protein F9D16.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05589
R:Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15419
A:Accession: T05589
A:Molecule type: DNA
A:Residues: 1-350 <BEV>
A:Cross-references: UNIPROT:Q9SUR9; EMBL:AL035394
A:Experimental source: cultivar Columbia; BAC clone F9D16
C:Genetics:
A:Map position: 4
A:Introns: 54/1; 77/2; 119/1; 153/2; 185/3; 207/3; 254/1; 280/3; 315/3
A:Note: F9D16.40

Query Match 60.9%; Score 1134; DB 2; Length 350;
Best Local Similarity 61.4%; Pred. No. 1.7e-69;
Matches 221; Conservative 56; Mismatches 69; Indels 14; Gaps 4;

Qy 3 ASDLESKAKAFVDDDFELATELYSQADAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db 2 AKELADKAKAFVDDDFVAVDLYSKAIDLDPNCABFFADRAQAQYIKLESGFTEAVADANK 61
Qy 63 AIELDPMMHKAYYKRGACIKLEBYQYAKALELGSYASGDSRPFARLLKCEDERIAES 122
Db 62 AIELDPSLTAKYLRKGTACMKLEBYRTAKTALERKASITPSESFKPLIDECNPLITEE 121
Qy 123 SQAPKVNVEATVAATIEDKEDFTNMENPPVIEPPSKP--KYRHDYNSATEVVLTIIFAK 180
Db 122 KDL-VQVPSTLPSV-----TAPPVSELDVTPTAKYRHEYQKPEEVVTVFAK 170
Qy 181 GVPADSVVDGEGQMLSVSIEVPGEYPHFOPRLFSKIIPKCKYQVLSTKVEIRLAKAE 240
Db 171 GIPKQNVNIDFGEQILSVSIEVPGEDAYLQPLRFGKIIPDKCYEVLSTKIEICLAKAD 230
Qy 241 QVTWTLIDYGRPKTVQKISTPAETAPRPSYPSSKAKQWDKLEAEVKKEEKLGDG 300
Db 231 IITWASLEHGKPAVLP-KPNVSSEVSORPAYPSKKVQWDKLEAEVKKEEKLGDG 289
Qy 301 AALNKFRRDIYKDAEDMRRAMKSFVESNGTSLSTNNKQVGAKKVGGSPDGMELKKWE 360
Db 290 AALNKFRRDIYQNAEDMRRAMKSFVESNGTSLSTNNKQVGTIESTPTDGMELKKWE 349

RESULT 3
T20305
hypothetical protein D1054.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20305
R:Matthews, P.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19255
A:Accession: T20305
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-198 <WILL>
A:Cross-references: UNIPROT:Q18949; EMBL:Z74030; PIDN:CAA98442.1; GSPDB:GN00023; CESP:D1
A:Experimental source: clone D1054
C:Genetics:
A:Gene: CESP:D1054.3
A:Map position: 5
A:Introns: 6/2; 71/3; 164/2

Query Match 20.6%; Score 383; DB 2; Length 198;
Best Local Similarity 41.7%; Pred. No. 5.8e-19;
Matches 86; Conservative 38; Mismatches 68; Indels 14; Gaps 6;

Qy 158 SKPKYRHDYNSATEVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEYPHFQPLRFSK 217
```

```
Db 3 SKP--RHDWFQSETDVTLTILKRGVPLDDCSVSLSD---NNTLTVRQCDEILFYGLSQG 57
Qy 218 IPEKCKYQVLSTKVEIRLAK-AEQVTTWTL--DYSGRPKTVQKISTPAETAPRPSYPS 274
Db 58 VKKDDLTVKTCTAAKVRVLPKFAERNRWASLLKDGQGVAAV--QSVSPNPEA-----PT 111
Qy 275 SKAKDWDKLEAEVKKEEKEELKDGDAALNKFFRDIYKDAEDMRRAMKSFVESNGTSL 334
Db 112 TTVKKNWDALEKQAVKEEDESLEGDAAVNKKFKMYNDASDDVRRAMKSYSESNGTSL 171
Qy 335 STNWKDVGAKKVGGSPDGMELKKWE 360
Db 172 STNWSBIGQKTECQPPACMEYKEYE 197

RESULT 4
S66940
SGT1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O2805; protein YOR057W
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C:Accession: S66940
R:Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66929
A:Accession: S66940
A:Molecule type: DNA
A:Residues: 1-395 <BOH>
A:Cross-references: UNIPROT:Q08446; EMBL:Z74965; NID:g1420194; PID:e251976; PID:g1420195
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:SGT1; GSS1; MIPS:YOR057W
A:Cross-references: SGD:S0005583; MIPS:YOR057W
A:Map position: 15R
A:Function:
A:Description: suppressor of G2 allele of SKP1

Query Match 19.6%; Score 364.5; DB 2; Length 395;
Best Local Similarity 29.3%; Pred. No. 2.6e-17;
Matches 108; Conservative 69; Mismatches 133; Indels 59; Gaps 14;

Qy 16 DDFELATELYSQADAGPATADLYADRAQAHIKLGNYTEAVADAN-KATELDPMMHKAY 74
Db 61 DATWNAKELLKAL----MTAEGRGDRSK--IGLVNFRYFVHFNIKDYEL----AQSY 110
Qy 75 YRGAACIKLEEQYQAKALELGSYASGDSRPFARLLKCEDERIAESQAPKVNVEATV 134
Db 111 FK-----AKNLGYVDDTLPWEDRLKLNKKKKQKQSTNKHITKPEV--- 155
Qy 135 AATIEDKEDFTNMEN--TPPVIE--PPSKYRHDYNSATEVLTIFAKGVPADSVVID 190
Db 156 --SIENRGDNSSHSPIPLKIETAPQESPKIDWYQSSTSVTISLFTVNLPSKEQVN 213
Qy 191 F-----GEQMLSVSIEVP--GEYPHFQPLRFSKIPEKCKYQVLSTKVEIRLAKAQVT 243
Db 214 IYISPNDRRTLSISYQVPKSGSE-FOYNAKLSHEVDPAVSLKIFPPKLEITLSKIDSTQ 272
Qy 244 WTTLT-----DYSGRPKTVQKISTPAETAPRPSYPSSKAKK-DWDKLEAEVKK 290
Db 273 WKLEEDILTSSRLSDEGKNDSATLLSAETASKERLSYPSSSKKKIDWKLIDIEEA 332
Qy 291 EEKKEKLDGDAALNKFFRDIYKDAEDMRRAMKSFVESNGTSLSTNNKQVGAKKVGGSP 350
Db 333 DEEAGSAD-----SFFQKLYAGADPDTKRAMKSFIESNGTALSTDWEDVSKGTVTSP 386
Qy 351 PDGMELKKW 359
Db 387 PEGMEPKHW 395

RESULT 5
T40307
```

hypothetical protein SPBC36.12c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40307
R;Lyne, M.; Wood, V.; Rajadream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh
submitted to the EMBL Data Library, May 1998
A;Reference number: 221919
A;Accession: T40307
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-444 <LYN>
A;Cross-references: EMBL:AL023589; PIDN:CAA19060.1; GSPDB:GN00067; SPDB:SPBC36.12c
A;Experimental source: strain 972h-; cosmid c36
C;Genetics:
A;Gene: SPDB:SPBC36.12c
A;Map position: 2
A;Introns: 151/2; 282/3

Query Match 18.3%; Score 341; DB 2; Length 444;
Best Local Similarity 31.3%; Pred. No. 1.2e-15;
Matches 98; Conservative 53; Mismatches 114; Indels 48; Gaps 11;

Qy 77 KGAACIKLEEQYAKAAELGSSVASGDS-----RPARLLKRCDCERIAESS-----QA 125
Db 150 RGLAYLLEQYQSSAICFGFALKKEDLTQWQVQDLSMLTIVHEQONQDTSSLIPDEC 209
Qy 126 PVKNVEATVAATIEDKEDFTNMENTPP-----VIEPPSKPKYRHDYNSATEVVLITPAK 180
Db 210 P--NIPELEAAKIEGDEFLNLSKAPQEQIEKNEKLSNRIRYDWSQTSPLNIDYAK 267
Qy 181 GVPADSVVDIGEOMLSVSVIEVPEEPVHFQ---PRLFSKILPEKCKYQVLSKVEIRL- 236
Db 268 KVKDEDVSLMEKNTLKIEIKL--EDGSIFSLVDPLYEEIVPEKSFKLFSSKVEITLI 325
Qy 237 AKABQVTTTLDYSGRPKTPQKISTPAETA---PRSPYPSKA-----KKDWDKLEA 286
Db 326 KKVSEIKWEAL-----VKSPANVNVYAKDSNHSASGNTKNAKADWDLSAK 373
Qy 287 EVKKEEKEKLDGDAALNKKFRDIYKDAEDMRAMKMSFVESNGTVLSTNWKDVGAKKV 346
Db 374 LADLEEDP--TGEAALANLFQNLKNADDTTRAMKMSYTESNGTALSTNWKDVKSKTF 431
Qy 347 EGSPDPGHELKW 359
Db 432 ETKPPQMEPKKF 444

RESULT 6
T48783
related to SGT1 protein [imported] - Neurospora crassa
N;Alternate names: protein 13E11.310
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: T48783
R;Schulze, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Reference number: 224541
A;Accession: T48783
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-469 <SCH>
A;Cross-references: UNIPROT:Q9P6Y5; EMBL:AL353820; GSPDB:GN00112; NCSP:13E11.310
A;Experimental source: cosmid contig 13E11; strain 74
C;Genetics:
A;Gene: NCSP:13E11.310
A;Map position: 2
A;Introns: 277/3

Query Match 17.6%; Score 327; DB 2; Length 469;
Best Local Similarity 24.1%; Pred. No. 1.1e-14;
Matches 117; Conservative 71; Mismatches 155; Indels 146; Gaps 16;

Qy 1 MAADLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60

Db 1 MSATTIAQKIDALAKGDHATAITNLDKALESSNSPAWLLA-RSKAQKAKNLEALHDA 59
Qy 61 NKAIELDPMHKAAYRKAACIKLE-----EYQTKAAALELGSSVASGD--SRFARLLKE- 113
Db 60 ELA-----YHAAERGSQTSRSHMTQAQYRRAVIYVQLG-REPADADCCAKWSMLLAEG 111
Qy 114 -----CDERIAESSQAPVKNVEATVAATIED----- 140
Db 112 RPAREDGVEKKVDSEGNVYTYDFELADKENQPKPTSDGNALAAAEKGPYSTDWNNAF 171
Qy 141 -----KEDFTNM-----ENTPPVIEP 156
Db 172 SWSQALGRLLKLPKDPHGWKVNVTKIPPRPEKKAKSPAEASASEDELKEQRPQKEA 231
Qy 157 PS-----KPKYRHDYNSATEVVLITFAKGVPADSVVIDFG-QLMSVSVIE-VPGSEP 207
Db 232 PPGSVSDERKMLRIDFYQNTQTVSLFVKDKVKEDLVKVEFGKQRTKVRISPIPREAA 291
Qy 208 YHFQPR-----LFSKILPEKCKYQVLSKVEIRLAKA----- 239
Db 292 PVKPGDRQATSTLVLAGEIDPSASRWSASPRKIELVLQKATPGVKWGRMGEEKIGIVES 351
Qy 240 -----EQVTTTLDYSGRPKTPQKISTPAETAPRPSYPSS--KAKDWDKLEAEVKKEEK 293
Db 352 DQEPATITATSSNAATAKALPSTSTPAKV---PAYTSSKSGPKAWDSLPVDDKEDG 408
Qy 294 BEKLDGDAALNKKFRDIYKDAEDMRAMKMSFVESNGTVLSTNWKDVGAKKVSGSPDG 353
Db 409 QD-----INGFPKTLTKGSTPEQORAMKMSFLESNGTTLSTNWDKVDKVPVTPDEG 461
Qy 354 MELKKW 359
Db 462 VEPKPW 467

RESULT 7
E84858
phosphoprotein phosphatase (EC 3.1.3.16) At2g42810 [similarity] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 28-Apr-2003
C;Accession: E84858
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-533 <STO>
A;Cross-references: GB:AE002093; NID:94512673; PIDN:AAD21727.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g42810
A;Map position: 2
A;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos
C;Keywords: phosphoric monoester hydrolase

Query Match 12.0%; Score 223; DB 2; Length 533;
Best Local Similarity 28.0%; Pred. No. 1.5e-07;
Matches 74; Conservative 43; Mismatches 99; Indels 48; Gaps 9;

Qy 3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db 13 ABEFKSQANEPKCHKYSSAIDLYTKAELNSNNAVYANPAFAHTKLEEYGSALQDASK 72
Qy 63 AIELDPMHKAAYRKAACIKLEEQYAKAAELGSSVASGDSRFAHLLKEDBRIA--- 119
Db 73 AIEVDYSYKGYRGAAYLAAMGFKDALKDFQVKRLSPNDPDATRKLECEKAVMKLK 132
Qy 120 -EESQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYNSATEVLTIF 178
Db 120 -EESQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYNSATEVLTIF 178

Db 133 PEEAISVPSE-RRSVASI-----DFPHIGNKPRSSMPTKTALA-----AVVAAMVVA 182
 Qy 179 AKGVADSVVDIFGQMLSVSI-----EVGEEPYHFQFRL-----FSKIIPKCK 223
 Db 183 VRGPATTEIL-----MVLVSVLGTFWWEV---EPQYSGARIEEGEVTLDVFKTMEDF 233
 Qy 224 K-----YQVLSTKVEIRLA 237
 Db 234 KNQKTLHKRYAYQIVLQTRQILLA 257

RESULT 8
 H86257
 protein F5011.2 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C;Accession: H86257
 R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: H86257
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-572 <STO>
 A;Cross-references: UNIPROT:Q9LNB6; GB:AE005172; NID:98778620; PIDN:AAF79628.1; GSPDB:GN
 C;Genetics:
 A;Gene: F5011.2
 A;Map position: 1

Query Match 10.4%; Score 193.5; DB 2; Length 572;
 Best Local Similarity 24.1%; Pred. No. 1.6e-05;
 Matches 82; Conservative 50; Mismatches 121; Indels 87; Gaps 13;

Qy 3 ASDLESKAKAFVDDDDPELATYLSQAIDAGPATADYADRAQAHIKLGNVTEAVADANK 62
 Db 2 ABEAKAGNAAPSSGDFTTAHNHTEATALAPTNRHVLFSNRSAAHLSHQAEALSDAKE 61
 Qy 63 AIELDPMHMKAYRKGAACIKLEEQYAKAALGLSSVAGSDRSFARLLKCEDERIAEES 122
 Db 62 TIKLPYWPKGYSRLGAHLGNQFELAVTAYKKG-----LDVDPITNEA 105
 Qy 123 SQAPVKVNEATVAAT-----IEDKEDFTNMNTPPV-----IEPPSKPKYRHYDYS 169
 Db 106 LKSLGDAEASVARSRAAPNPFDAFGPENMTWTKLTSDPSTRGFLQP-----DFVNM 158
 Qy 170 ATEV-----VLTIFAKGVPADVVIDFGQMLSVSIEVPGEEPHFQFRLFSKIIPKCK 224
 Db 159 MQELQKQPSLNLVYK-----DQVW-----QSLGVLLNVKFRPP-----PPQGEAEVPPSDM 207
 Qy 225 YQVLSTKVEIRLAKAEQVTTTLDYSGRPKTVPOKISTPAETAPRPSYPSSKAKQWMDKL 284
 Db 208 QQSSSNEPEV-----EKKREP-EPEPEPEVEEKEKE-----239
 Qy 285 EAEVKEEKE-----EKLGDGAALNKFRIYKDAEDM 318
 Db 240 RKEAKKEKEGLNAAYKKKDPETAIOHYSTAIREID-DEDI 278

RESULT 9
 S52570
 phosphoprotein phosphatase (EC 3.1.3.16) 5 [validated] - human
 N;Alternate names: serine/threonine phosphatase pp5
 C;Species: Homo sapiens (man)
 C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C;Accession: S52570; PC4136

R;Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.
 EMBO J. 13, 4278-4290, 1994
 A;Title: A novel human protein serine/threonine phosphatase, which possesses four tetra
 A;Reference number: S52570; MUID:95009929; PMID:795273
 A;Accession: S52570
 A;Molecule type: mRNA
 A;Residues: 7-498 <CH2>
 A;Cross-references: UNIPROT:P53041; EMBL:S73586; EMBL:X89416
 A;Experimental source: teratocarcinoma cell line NTERA-2
 R;Xu, X.L.; Lagercrantz, J.; Zickert, P.; Bajalica-Lagercrantz, S.; Zetterberg, A.
 Biochem. Biophys. Res. Commun. 218, 514-517, 1996
 A;Title: Chromosomal localization and 5' sequence of the human protein serine/threonine
 A;Reference number: PC4136; MUID:96144708; PMID:8561788
 A;Accession: PC4136
 A;Molecule type: mRNA
 A;Residues: 1-37 <XUX>
 A;Cross-references: EMBL:X92121; NID:gl1177477; PIDN:CAA63089.1; PID:e205526; PID:gl117747
 A;Experimental source: fetal brain
 C;Comment: mRNA encoding this protein was detected in all human tissues tested and was 1
 C;Genetics:
 A;Gene: GDB:PPP5C; PPP5; PP5
 A;Cross-references: GDB:I36857
 A;Map position: 19q13.3-19q13.3
 C;Function:
 A;Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine [vali
 A;Note: may play a role in the regulation of RNA synthesis and mitosis
 C;Superfamily: protein phosphatase 5; phosphoesterase core homology
 C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
 F;28-61/Domain: tetra-trico-peptide repeat homology <TT1>
 F;62-95/Domain: tetra-trico-peptide repeat homology <TT2>
 F;96-125/Domain: tetra-trico-peptide repeat homology <TT3>
 F;204-467/Domain: phosphoprotein phosphatase core homology <PPP>
 F;236-305/Domain: phosphoprotein phosphatase core homology <PPC>
 F;242,244,271/Binding site: iron (Asp, His, Asp) #status predicted
 F;271,303,352,426/Binding site: zinc (Asp, Asn, His, His) #status predicted
 F;274,304,450/Active site: Asp, His, Tyr #status predicted
 F;275,399/Binding site: substrate phosphate (Arg) #status predicted

Query Match 10.3%; Score 192; DB 1; Length 498;
 Best Local Similarity 24.7%; Pred. No. 1.7e-05;
 Matches 62; Conservative 47; Mismatches 92; Indels 50; Gaps 6;

Qy 3 ASDLESKAKAFVDDDDPELATYLSQAIDAGPATADYADRAQAHIKLGNVTEAVADANK 62
 Db 28 AEELKTOANDYFRKADYENAIKFYSQAIELNPNAIYGNRSLAYLRETCYGYALGDATR 87
 Qy 63 AIELDPMHMKAYRKGAACIKLEEQYAKAALGLSSVAGSDRSFARLLKCEDERIAEES 122
 Db 88 AIELDKKIKYGYRRRAASNNALGKFRALRDYETTVKVKPHDKDAKMKYQECNKIVKQKA 147
 Qy 123 SQAPVKVNEATVAATIEDKEDFTNMNTPPVIEPPSKPKYRHYDYSNATEVLTIFAKGV 182
 Db 148 FERAIAGDEH--KRSVVDSLIDSM---TIEDEYSGPK-----L 181
 Qy 183 PADSVVIDFGQMLSVSIEVPGEEPHFQFRLFSKIIPKCKYQV-----LSTKV 232
 Db 182 EDGKVTISFMKELM-----QWYKQDKKLH-----RKCAVQILVQVKEVLSKSLTLV 227
 Qy 233 EIRLAKAEQVT 243
 Db 228 ETTIKETEKIT 238

RESULT 10
 T00954
 hypothetical protein F20D22.4 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
 C;Accession: T00954
 R;Vysotskaya, V.S.; Osborne, B.I.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji,
 K.; Peng, J.; Kim, C.; Kurtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.;
 submitted to the EMBL Data Library, May 1998
 A;Description: Arabidopsis thaliana chromosome 1 BAC F20D22 complete sequence.

A;Residues: 1-589 <NIC>
A;/Cross-references: UNIPROT:P15705; GB:M28486; NID:g172765; PIDN:AAA35121.1; PID:g172766
R:/de Haan, M.; Maarse, A.C.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995

A;Reference number: S54617
A;Accession: S54633
A/Molecule type: DNA

A;Residues: 1-589 <DEH>
A;/Cross-references: EMBL:X87331; NID:g1041652; PID:g829137
R:/de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996

A;Reference number: S66877
A;Accession: S66893
A/Molecule type: DNA

A;Residues: 1-589 <DEM>
A;/Cross-references: EMBL:Z74935; NID:g1420138; PID:e252326; PID:g1420139; MIPS:YOR027w
A;/Experimental source: strain S288C

C;Genetics:
A/Gene: SGD:STI1

A;/Cross-references: SGD:S000553; MIPS:YOR027w
A;/Map position: 15R

C/Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat homology F/5-38/Domain: tetratricopeptide repeat homolog <TT1>

F/40-73/Domain: tetratricopeptide repeat homolog <TT2>

F/74-107/Domain: tetratricopeptide repeat homolog <TT3>

Query Match 9.5%; Score 177.5; DB 2; Length 589;
Best Local Similarity 33.3%; Pred. NO. 0.0002;
Matches 52; Conservative 22; Mismatches 59; Indels 23; Gaps 4;

QY 3 ASDLESKAKAEAFVDDDFELATPELYSOADTACGAPATADLVADRQAQHKLGNYTEAVADVANK 62
Db : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :
396 AEARLEGGKYFFTKSDPNVAKVATEWKRAPEDARGYSNRRAALAKLMSPFEAIADCNK 455
QY 63 AIELDPMMHKAYRYKGAACTIKLEYQTAKAAALE-----GSS-----YASGD 104
Db |||||:::||:||||:|:
456 AIEKDPNFVRAYIRKAQIAVKAYSALAETLDARTDAEVNGSGSAREIDIQLYYKASQ 515
QY 105 SFPAULLKECDERTAESOAFPKNVEATVAATTIED 140
Db |||||---PGTSNETPETYGRAMKDPE--VAAIMOD 546

RESULT 15
S56658 stress-induced protein stil - soybean
C;Species: Glycine max (soybean)
C;/Date: 08-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 16-Aug-2004
C;/Accession: S56658; S56659; S71763
R;/Torres, J.H.; Chatteillard, P.; Stutz, E.
Plant Mol. Biol. 27, 1221-1226, 1995

A;/Title: Isolation and characterization of gmsti, a stress-inducible gene from soybean (<C;Genetic Engineering Laboratory>)

A;/Reference number: S56658; PMID:7766904
A;Accession: S56658
A;/Status: nucleic acid sequence not shown

A;/Molecule type: mRNA

A;Residues: 1-569 <TOR>
A;/Cross-references: UNIPROT:Q43468; EMBL:X79770
A;Accession: S56659
A;/Molecule type: DNA

A;Residues: 417-569 <TOW>
A;/Cross-references: EMBL:X79770
R;/Stutz, E.
submitted to the EMBL Data Library, June 1994

A;/Reference number: S71763
A;Accession: S71763
A;/Molecule type: mRNA

A;Residues: 1-24, F., 26-569 <STU>
A;/Cross-references: EMBL:X79770; NID:g872115; PID:g872116
C;Genetics:
A/Gene: stl1

A;/Introns: 491/2; 513/3; 533/3
A;/Note: The list of introns is incomplete

C;Superfamily: tetratricopeptide repeat homology
F:321-354/Domain: tetratricopeptide repeat homology <TT1>
E:381-414/Domain: tetratricopeptide repeat homology <TT2>
F:415-448/Domain: tetratricopeptide repeat homology <TT3>
F:449-482/Domain: tetratricopeptide repeat homology <TT4>

Query Match 9.5%; Score 176; DB 2; Length 569;
Best Local Similarity 33.3%; Pred. No. 0.00024;
Matches 40; Conservative 17; Mismatches 63; Indels 0; Gaps 0;
Qy 3 ASDLESKAKEAFVDDDFELATELYSQATDAGPATADLYADRAQAHIKIGNYTEAVADANK 62
Db 381 ADEAREKGNELFKQKYPEATKHYTEALKRNPDKAKAYSNRAACVTKLGAMPEGLKDAEK 440
Qy 63 AIELDPMMHKAYYRKGACIKLEIYQTKAAALELGSSYASGDSRFARLLKECDERIAEES 122
Db 441 CIELDPTFSKGYTRKGAQVQFSMKEYDKALETYREGLKHPNPNQELLDGIRRCVQINKAS 500

Search completed: April 20, 2005, 18:35:07
Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:17:45 ; Search time 178 Seconds
(without alignments)
1038.543 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESKAKEAFVDDDFE.....GAKKVEGSPDGMELKKWEY 361

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1594.5	85.7	373	2 Q8W516	Q8W516 hordeum vul
2	1592.5	85.6	367	2 Q9FPB1	Q9FPB1 oryza sativ
3	1589.5	85.4	367	2 Q9SE32	Q9SE32 oryza sativ
4	1243	66.8	370	2 Q84UV7	Q84UV7 nicotiana b
5	1142	61.4	350	2 Q8W515	Q8W515 arabidopsis
6	1138	61.1	354	2 Q704T1	Q704T1 brassica ol
7	1135	61.0	358	2 Q9SUT5	Q9SUT5 arabidopsis
8	1134	60.9	350	2 Q9SUT9	Q9SUT9 arabidopsis
9	1133.5	60.9	327	2 Q9SM53	Q9SM53 rubus idaeu
10	1125.5	60.5	355	2 Q704T2	Q704T2 brassica ol
11	870	46.7	273	2 Q84LL4	Q84LL4 arabidopsis
12	615	33.0	333	1 SUGT HUMAN	Q9Y2Z0 homo sapien
13	608.5	32.7	331	2 Q6IP01	Q6IP01 xenopus lae
14	602.5	32.4	211	2 Q8LKE8	Q8LKE8 nicotiana b
15	602.5	32.4	211	2 Q8LKE9	Q8LKE9 nicotiana b
16	599.5	32.2	336	1 SUGT MOUSE	Q9CX34 mus musculu
17	594	31.9	365	2 Q6VX76	Q6VX76 homo sapien
18	423	22.7	375	2 Q6BXK8	Q6BXK8 debaryomyce
19	383	20.6	198	2 Q18949	Q18949 caenorhabdi
20	370.5	19.9	377	2 Q6FKN1	Q6FKN1 candida gla
21	364.5	19.6	395	1 SGT1 YEAST	Q08446 saccharomyc
22	364.5	19.6	395	2 Q6Q514	Q6Q514 saccharomyc
23	348.5	18.7	379	2 Q6CJ85	Q6CJ85 kluyveromyc
24	341	18.3	379	1 G1T7 SCHPO	Q59709 schizosacch
25	334.5	18.0	466	2 Q3P675	Q3P675 neurospora
26	329	17.7	221	2 Q9N8L0	Q9N8L0 trypanosoma
27	321.5	17.3	204	2 Q7Q5W9	Q7Q5W9 anopheles g
28	317	17.0	178	2 Q9VHT3	Q9VHT3 drosophila
29	314	16.9	178	2 Q8SY87	Q8SY87 drosophila
30	313.5	16.8	247	2 Q7TQJ2	Q7TQJ2 rattus norv
31	306.5	16.5	462	2 Q7RV15	Q7RV15 neurospora

ALIGNMENTS

RESULT 1

ID	Q8W516	PRELIMINARY;	PRT;	373 AA.
AC	Q8W516;	AC		
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	SGT1.			
GN	Name=SGT1;			
OS	Hordeum vulgare (Barley).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;			
OC	Triticeae; Hordeum.			
OX	NCBI_TaxID=4513;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21893744; PubMed=11847307; DOI=10.1126/science.1067554;			
RA	Azevedo C., Sadanandom A., Kitagawa K., Freialdenhoven A., Shirasu K.,			
RA	Schulze-Lefert P.;			
RT	"The RAR1 interactor SGT1, an essential component of R gene-triggered			
RT	disease resistance.";			
RL	Science 295:2073-2076(2002).			
DR	EMBL; AF439974; AAL33610.1; -.			
DR	Pfam; PF04969; CS; 1.			
DR	Pfam; PF05002; SGS; 1.			
DR	Pfam; PF00515; TPR_1; 1.			
DR	SMART; SM00028; TPR; 3.			
DR	PROSITE; PS50005; TPR; 1.			
DR	PROSITE; PS50293; TPR_REGION; 1.			
KW	Repeat; TPR repeat.			
SQ	SEQUENCE 373 AA; 40997 MW; 74A6E2C77D0F9EAD CRC64;			
Query Match 85.7%; Score 1594.5; DB 2; Length 373;				
Best Local Similarity 83.0%; Pred. No. 1e-95;				
Matches 307; Conservative 26; Mismatches 26; Indels 11; Gaps 2;				
Qy	2 AASDLESKAKEAFVDDDFELATELYSQADAGPATADLYADRAQAHIKLGNYTEAVADAN	61		
Db	5 AASDLESKAKEAFVDDDFELAAELYTQAIIEAGPATAEYADRAQAHIKLGSYTEAVADAN	64		
Qy	62 KATLEDPMHKAYVRKGAACIKLEEYOTAKAALGSSYASGDSRFAFLKKECDERTAAE	121		
Db	65 KATLEDPMSHKAYLRKGSACIKLEEYOTAKAALGVSSYASGDSRFRFLMKCCDDRTAAE	124		
Qy	122 SSQAPVKNVEATVA-----ATTEDKEDFTNMENTPPVIPPSPKPKYRHDYINSAT	171		
Db	125 ASQAPVKNAAAVAPATSSGATTVTTEADQDQDGENMENAQPTVEVSPKPKYRHDYINPT	184		
Qy	172 EVLVTIFAKGVADSVVDVDFGEQMLSVSIEVPGEYHFQRLFSKIIPKCKYQVLSSTK	231		
Db	185 EVLVTIFAKGVADSVVDVDFGEQMLSVSIEIPGEYHFQRLFSKIIPKCKYQVLSSTK	244		
Qy	232 VETRLAKAEQVTTTLDYSGRPKTPVKISTPATETAPRPSYPSSKAKDWDKLEAEVKKE	291		

```

Db 245 VEIRLAKAEPTVWTSYDYGKPK-APQKINVAESAQRPSYPSSKSKDWDKLEAEVKKQ 303
Qy 292 EKEBKLGDGALNKFPRDIYKDAEDMRRAMKSFVESNGTVLSTNWKDYGAKKVEGSP 351
Db 304 EKDEKLGDGALNKFPRDIYKDAEDMRRAMKSFVESNGTVLSTNWKDYGAKKVEGSP 363
Qy 352 DGMELKKWEY 361
Db 364 DGMELKKWEY 373

RESULT 2
Q9FPB1
ID Q9FPB1 PRELIMINARY; PRT; 367 AA.
AC Q9FPB1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sgt1.
GN Name=P0006C01.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hanada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saij S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
DR EMBL; AP002744; BAB19060.1; -
DR Gramene; Q9FPB1; -
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR007699; SGS.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR Pfam; PF0515; TPR; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00005; TPR; 1.
DR PROSITE; PS0293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 367 AA; 40938 MW; 885E7758A754FED1 CRC64;

Query Match 85.6%; Score 1592.5; DB 2; Length 367;
Best Local Similarity 84.1%; Pred. No. 1.4e-95;
Matches 307; Conservative 27; Mismatches 24; Indels 7; Gaps 2;

Qy 2 AASDLESKAEAFVDDDFELATLAYSQAI DAGPATADLYADRAQAHIKLGNYTEAVADAN 61
Db 5 AASDLESKAEAFVDDDFELATLAYSQAI DAGPATADLYADRAQAHIKLGNYTEAVADAN 64
Qy 62 KAIELDPMMHKAYRKGAACIKLEEYQTAKAALBELGSSYASGDSRFRLLKCEDERIAEE 121
Db 65 KAIELDPMMHKAYRKGAACIKLEEYQTAKAALBELGSSYASGDSRFRLLKCEDERIAEE 124
Qy 122 SSQAPVKNE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYNSATEVLT 176

us-10-609-078-8.rup
Db 125 LSEVPVKAEDGAAAPSVASFEKDDAANMDNTPMVE--VKPKYRHDYNSATEVLT 182
Qy 177 IFAGKVPADSVVLDGEGOMLSVIEVPGEPYHFQRLFSKIIPEKCKYQVLTSTKVEIRL 236
Db 183 IFAGKVPADSVVLDGEGOMLSVIEVPGEPYHFQRLFSKIIPEKCKYQVLTSTKVEIRL 242
Qy 237 AKAEQVTTWTLTDYSGRPKTVQPKISTPAETAPRPSYSSKAKKDWKLEAEVKKEEKEK 296
Db 243 AKAEQITWTSYDYGKPKAVPQKIIPAESAQRPSPSSKSKDWDKLEAEVKKEEKEK 302
Qy 297 LGDGAALNKFPRDIYKDAEDMRRAMKSFVESNGTVLSTNWKDYGAKKVEGSP 356
Db 303 LEGDGAALNKFPRDIYKDAEDMRRAMKSFVESNGTVLSTNWKDYGAKKVEGSP 362
Qy 357 KWWEY 361
Db 363 KWWEY 367

RESULT 3
Q9SE32
ID Q9SE32 PRELIMINARY; PRT; 367 AA.
AC Q9SE32
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sgt1.
GN Name=Sgt1;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20035956; PubMed=10571178; DOI=10.1016/S0092-8674(00)81522-6;
RA Shirasu K., Lahaye T., Tan M.W., Zhou F., Azevedo C.,
RA Schulze-Lefert P.;
RT "A novel class of eukaryotic zinc-binding proteins is required for
RT disease resistance signaling in barley and development in C.
RT elegans."
RL Cell 99:355-366(1999).
DR EMBL; AF192467; AAF18438.1; -
DR Gramene; Q9SE32; -
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR007699; SGS.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR Pfam; PF0515; TPR; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00005; TPR; 1.
DR PROSITE; PS0293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 367 AA; 40952 MW; 885F320DF203FED1 CRC64;

Query Match 85.4%; Score 1589.5; DB 2; Length 367;
Best Local Similarity 83.8%; Pred. No. 2.1e-95;
Matches 306; Conservative 28; Mismatches 24; Indels 7; Gaps 2;

Qy 2 AASDLESKAEAFVDDDFELATLAYSQAI DAGPATADLYADRAQAHIKLGNYTEAVADAN 61
Db 5 AASDLESKAEAFVDDDFELATLAYSQAI DAGPATADLYADRAQAHIKLGNYTEAVADAN 64
Qy 62 KAIELDPMMHKAYRKGAACIKLEEYQTAKAALBELGSSYASGDSRFRLLKCEDERIAEE 121
Db 65 KAIELDPMMHKAYRKGAACIKLEEYQTAKAALBELGSSYASGDSRFRLLKCEDERIAEE 124
Qy 122 SSQAPVKNE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYNSATEVLT 176

```



```

Db 125 LTEVPVKAEDGAAAPS VASFEVEEKDDAANMDNTPPMVE--VKPKYRHDFYNSATEVVLT 182
Qy 177 IFAGVPADSVVDVDFGEOMLSVSEVPGEEYHFPQPLFSKIIPEKCKYQVLSTKVIRL 236
Db 183 IFAGVPAENVVVDVDFGEOMLSVSEVPGEEYHFPQPLFSKIIPEKGRYQVLSTKVIRL 242
Qy 237 AKAEQVTTWTLTLDYSGRPKTPVKISTPAETAAPRPSYSSKAKDWDKLEAEVKKEEK 296
Db 243 AKAEQITWTLTLDYDKPKAVPKIIPPAESQRPSYSSKSKDWDKLEAEVKKEEK 302
Qy 297 LDGDAALNKFPRDIYKDAEDMRAMKSFVESNGTVLSTNWKDVGAKKVGGSPDGMEL 356
Db 303 LEGDAALNKFPRDIYSDADEMRAMKSFVESNGTVLSTNWKDVGGKVGSPDGMEL 362
Qy 357 KKWEY 361
Db 363 KKWEY 367

RESULT 4
Q84UV7
ID Q84UV7 PRELIMINARY; PRT; 370 AA.
AC Q84UV7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SGT1a
OS Nicotiana benthamiana.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OC NCBI_TaxID=4100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22115173; PubMed=12119369;
RA Liu Y., Schiff M., Serino G., Deng X.W., Dinesh-Kumar S.P.;
RT "Role of SCP ubiquitin-ligase and the Cop9 signalosome in the N gene-
RT mediated resistance response to Tobacco mosaic virus.";
RL Plant Cell 14:1483-1496(2002).
DR EMBL; AF494083; AA085509.1; -.
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF00515; TPR_1; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 370 AA; 41240 MW; 88EF4C3408F5121F CRC64;

Query Match
Best Local Similarity 66.8%; Score 1243; DB 2; Length 370;
Matches 245; Conservative 48; Mismatches 63; Indels 14; Gaps 5;

Qy 3 ASDLESKAEAFVDDDFELATYLSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db 2 ASDLETRAKEAFIDDFELAVDLYTQALMTPKNAELFADRAQANIKLNYTEAVVADANK 61
Qy 63 AIELDPMHKYKYGKACIKLEYQTKAALELGSSVSGDSFARLLKCEDRIAEES 122
Db 62 AIELDPSMSKAYLRKGLACMKLEYQTKAALELGASLPAESRFTKLKCEDRIAEAE 121
Qy 123 SOAPVKVNEATVAATIED--KEDFTNMENPPVIEP-----PSKPKYRHDYNSAT 171
Db 122 GELPNQSVDKTSGNVVAPPASESLDNVAVPKDQPTVNLISYQGSAAKPRTRHFFYKPE 181
Qy 172 EVLITIFAKGVADSVVDVDFGEOMLSVSEVPGEEYHFPQPLFSKIIPEKCKYQVLSTK 231
Db 182 EVVVTIFAKGIPAKNVVDVDFGEQILSVSIDVPGDETYSFQPLFGKLTAKRYEVMSTK 241
Qy 232 VEIRLAKAEQVTTWTLTLDYSGRPKTPVKISTPAETAAPRPSYSSKAKK--DWDKLEAEVKK 290

```

```

Db 242 IEIRLAKAEPLHWTSLVET--RASAVQVRPNVSSD--APRPSYSSKLRHVDWDKLEAEVKK 299
Qy 291 EEKEEKLDGDAALNKFPRDIYKDAEDMRAMKSFVESNGTVLSTNWKDVGAKKVGGSP 350
Db 300 EEKDEKLDGDAALNKFPRDIYKDAEDTRRAMKSFVESNGTVLSTNWKVGTGKVEGSP 359
Qy 351 PDGMELKKWE 360
Db 360 PDGMELKKWE 369

RESULT 5
Q8W515
ID Q8W515 PRELIMINARY; PRT; 350 AA.
AC Q8W515
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SGT1a
GN Name=SGT1a;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21893744; PubMed=11847307; DOI=10.1126/science.1067554;
RA Azevedo C., Sadanandom A., Kitagawa K., Freialdenhoven A., Shitazu K.,
RA Schulze-Lefert P.;
RT "The RAR1 interactor SGT1, an essential component of R gene-triggered
RT disease resistance.";
RL Science 295:2073-2076(2002).
DR EMBL; AF439975; AAL33611.1; -.
DR HSP; P31948; 1ELW.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR SMART; SM00028; TPR_1; 2.
DR PROSITE; PS50005; TPR; 2.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 350 AA; 39277 MW; D23DBE990C817FB4 CRC64;

Query Match
Best Local Similarity 61.4%; Score 1142; DB 2; Length 350;
Matches 222; Conservative 56; Mismatches 68; Indels 14; Gaps 4;

Qy 3 ASDLESKAEAFVDDDFELATYLSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db 2 AKELADKAEAFVDDDFDVAVDLYSKAIDLDPNCAEFADRAQAYIKLESFTEAVADANK 61
Qy 63 AIELDPMHKYKYGKACIKLEYQTKAALELGSSVSGDSFARLLKCEDRIAEES 122
Db 62 AIELDPSLTAKYLRKGTACMKLEYRTAKTALALEGASITPSESFKKLIDECNFLTIEE 121
Qy 123 SOAPVKVNEATVAATIEDKEDFTNMENPPVIEPSPK--KYRHDYNSATEVVLTPAK 180
Db 122 KDL-VQVPSTLSPSV-----TAPVSELDTVTAKYRHEYQKPEVVVTPAK 170
Qy 181 GVPADSVVDVDFGEOMLSVSEVPGEEYHFPQPLFSKIIPEKCKYQVLSTKVIRLAKAE 240
Db 171 GIPKQNVNIDFGEQILSWIEVPGEDAYLQPLFGKIIIPDKCKYEVLSKIEIRLAKAD 230
Qy 241 QVTTWTLTLDYSGRPKTPVKISTPAETAAPRPSYSSKAKDWDKLEAEVKKEEKLDGD 300
Db 231 IITWASLEHGKGAFLP--KPNVSEVSQRPAVPSKVKQWDKLEAEVKKEKDEKLEGD 289
Qy 301 AALNKFPRDIYKDAEDMRAMKSFVESNGTVLSTNWKDVGAKKVGGSPDGMELKKWE 360
Db 290 AALNKFPRDIYQNAEDMRAMKSFVESNGTVLSTNWKVGTGKTIESTPPDGMELKKWE 349

```

```

RESULT 6
Q704T1 PRELIMINARY; PRT; 354 AA.
AC Q704T1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SGT1-like protein.
GN Name=sgt1b-2;
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Yemm A.A., Holub E., Tor M.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620883; CAF06581.1; -.
DR HSP; P53041; 1A17.
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR007699; SGS.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR Pfam; PF00515; TPR; 1; 2.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS0005; TPR; 1.
DR PROSITE; PS0293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 354 AA; B97D6F8A92285A97 CRC64;

Query Match 61.1%; Score 1138; DB 2; Length 354;
Best Local Similarity 62.4%; Pred. No. 4.8e-66;
Matches 224; Conservative 53; Mismatches 76; Indels 6; Gaps 4;

Qy 3 ASDLESKAKAFVDDDELATYLSQADTADYADRAQAHKLGNYTEAVADANK 62
Db 2 ASELAEKAKAFGLDDDFVADLYSKAIDLDPSCASFADRAQANIKLNFTEAVADANK 61
Qy 63 AIELDPMHXYKRGKACIKLEYYQYAKALELGGSSVAGDSRFAILLKCEDEIAEES 122
Db 62 AIELEPTLAKYLRKAGCKLEAYATAKALEGASVAPNESKFEKWDIECNLLIAEE 121
Qy 123 SQAPKVNVEATVAATIEDKEDFTNMNTPPVIEPPSKPKYRHDYNSATEVLTIFAKGV 182
Db 122 KDL-VQQVPPTLPS--STTPLAIAADAPPA--APAKPMFRHEFYQKPEEVGVAIFAKGI 176
Qy 183 PASVVDVDFEOMLSVIEVGPGEYPHFOPRLFSKIIPECKYQVLTSTKVEIRLAKAEV 242
Db 177 PKQNVNVEFGDQLSVVVDVAGEAYHFQPRLFQKIIPEKCRVEVLSTKVEIRLAKAEIV 236
Qy 243 TWTLDYSGRPKTPQKISTPAPRPSYSSKAKKWDKLEAEVKKKEEKLDGDAA 302
Db 237 TWASLEY-GKQALLPKNVASAVSQRPVTPSSKFGDWKLEAEVKKQEKDKLDGDAA 295
Qy 303 LNKFRRDIYKDAEDMERAMKSFVSNGTVLTSTNWKDVGAKKVGGSPPDGMELKKWEY 361
Db 296 MNKEFFSDIYQSAEDMERAMKSFVSNGTVLTSTNWKVEVGTGKVESTPPDGMELKKWEY 354

RESULT 7
Q9SUT5 PRELIMINARY; PRT; 358 AA.
AC Q9SUT5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein At4g11260 (SGT1a) (At4g11260_F8L21_50) (SGT1b)

```

```

DE SGT1-like protein).
GN Name=AT4g11260; Synonyms=At4g11260, SGT1b, sgt1a;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [5]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
MEDLINE=21893744; PubMed=11847307; DOI=10.1126/science.1067554;
RA Avevedo C., Sadanandom A., Kitagawa K., Freialdenhoven A., Shirasu K.,
RA Schuize-Lefert P.;
RT "The RARI interactor SGT1, an essential component of R gene-triggered
RT disease resistance.";
RL Science 295:2073-2076(2002).
RN [9]
RP SEQUENCE FROM N.A.
MEDLINE=22882917; PubMed=14504394; DOI=10.1073/pnas.2035132100;
RA Pawlowski J., Holzman M., Berney S.C., Fahrni J., Gooday A.J.,
RA Cedhagen T., Habura A., Bowser S.S.;
RT "The evolution of early Foraminifera.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11494-11498(2003).
RN [10]
RP SEQUENCE FROM N.A.
RA Shen W.H.;

```



```

Db      2 AKELADKAKEAFVDDDFVAVDLYSKAIIDLPNCAEFADRAQAQVIKLESTFEAVADANK 61
QY      63 AIELPMMHKAYRKGACIKLEEYQTAKAALGSSVAGSDRSFARLLKCEDERABES 122
Db      62 AIELDPSLTAKYLKRGATCKMKEEYRTAKTALEKASITPSEKFKKLIIDECNFIITEE 121
QY      123 SOAPVKNVEATVAATIEKEDFTNMENTPPVIEPPSKP--KYRHDYNSATEVVLITPAK 180
Db      122 KDL-VQPVSTLPSV-----TAPPVSELDVTPTAKYRHEYYQKPEEVVVVFAK 170
QY      181 GVPADSVVIDFGEOMLSVIEVPGEPYHFPORLSKLIPEKCKYQVLTSTKVEIRLAKAE 240
Db      171 GIPKQNVNIDFGEQLSVVIEVPGEDAYLQRLFGKLIIPDKCKYEVLTSTKIEICLAKAD 230
QY      241 QVTTWTLDSGRPKTVPOKISTPATPRPSPSSKAKKDWKLEAEVKKKEEKLGDG 300
Db      231 IITWASLEHGKGPVLP-KPNVSSEVSQRPAYPSSKVKYKDWKLEAEVKKKEEKLGDG 289
QY      301 AALNKFRRDIYKDAEDMRRAMKSFVBSNGTGLSTNNKDWGAKKVEGSPDGMELKKE 360
Db      290 AALNKFRRDIYQNADEDMRRAMKSFVBSNGTGLSTNNKQEVGTGTIESTPPDGMELKKE 349

RESULT 9
Q9SM53 PRELIMINARY; PRT; 327 AA.
AC Q9SM53
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein (fragment).
OS Rubus idaeus (Raspberry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Rosales; Rosaceae; Rosoideae; Rubus.
OX NCBI_TaxID=32247;
RN [1]
RP SEQUENCE FROM N.A.
RA Schroeder G., Lurz G., Schroeder J.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ251317; CAB61630.1; -
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR001440; TPR.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF00515; TPR; 1.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein.
FT NON-TER
SQ SEQUENCE 327 AA; 36617 MW; A7CB03A42D2032B8 CRC64;

Query Match 60.9%; Score 1133.5; DB 2; Length 327;
Best Local Similarity 66.2%; Pred. No. 8.5e-66;
Matches 223; Conservative 44; Mismatches 49; Indels 21; Gaps 5;

QY 34 PATADLYADRAQAHIKLGNTYEAVADANKAIELDPMHMKAYRKGACIKLEEYQTAKAA 93
Db 1 PKNSELYSDRAQANKSNLNTTEAVADANKAIELDPSLFKAYLRKGIACMKLEEVRTAKAA 60
QY 94 LEIGSSVAGSDRSFARLLKCEDERIAESSQAPVKNVEATVAATI-EDKEDFTNMENTPP 152
Db 61 LEGLASLAPGESFARKIMKCEDELIAEENGEPKQPMETTTTTEIABDVDP-----111
QY 153 VIEPPS-----KPKYRHDYNSATEVVLITPAKGVADSVVIDFGEOMLSVIEVPG 204
Db 112 -VDPSPNEVTAPVKPKYRHEFYQKAEVVVTFPAKIPAEVAVDPPGQILSVSDVPP 170
QY 205 EEPVHFPORLSKLIPEKCKYQVLTSTKVEIRLAKAEQVTTWTLDSGRPKTVPOKISTPA 264
Db 171 EVHIHFQRLFGKLIPEKCKRFEVLSTKVEIRLAKAEPNHNWTSLEFS-KDNPFVLKNAPV 229

```

```

QY 265 ETAPRPSPSSKAKK-DWMDKLEAEVKKKEEKLGDGAALNKFRRDIYKDAEDMRRAMM 323
Db 230 IGAQRSPSPSSKPRVDWMDKLEAQVKKKEKDEKLGDGAALNKFRRDIYKDAEDMRRAMR 289
QY 324 KSFVESNGTGLSTNNKDWGAKKVEGSPDGMELKKE 360
Db 290 KSFVESNGTGLSTNNKEVGNKVEGSPDGMEMKKE 326

RESULT 10
Q704T2 PRELIMINARY; PRT; 355 AA.
ID Q704T2
AC Q704T2;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE SGT1-like protein.
GN Name=sgt1b-1;
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RA Yemm A.A., Holub E., Tor M.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620882; CAF06580.1; -
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR007699; SGS.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR Pfam; PF00515; TPR; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50005; TPR; 2.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 355 AA; 39407 MW; BB4CAB6C407EB719 CRC64;

Query Match 60.5%; Score 1125.5; DB 2; Length 355;
Best Local Similarity 59.8%; Pred. No. 3.1e-65;
Matches 219; Conservative 57; Mismatches 71; Indels 19; Gaps 4;

QY 3 ASDLESKAKEAFVDDDFELATELYSOAIDAGPATADLYADRAQAHIKLGNTYEAVADANK 62
Db 2 ASELAEKAKEAFLEDDFDVAVDLYSRAIDLPNCAAFADRAQAQIKILNFTTEAVADANK 61
QY 63 AIELDPMHMKAYRKGACIKLEEYQTAKAALGSSVAGSDRSFARLLKCEDERIAES 122
Db 62 AIELEFSLKAYLRKGTACKMKEEYSTAKAALQKGASVAPNESKFNKLIIDECNLHIAEE 121
QY 123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVI-----EPPSKPKYRHDYNSATEVVL 175
Db 122 -----KDLAQOMPPL-----PSSSTTTPPLATAADSPAPAKPMERHEFYQKPEVVV 170
QY 176 TPAKGVADSVVIDFGEOMLSVIEVPGEPYHFPORLSKLIPEKCKYQVLTSTKVEIR 235
Db 171 TVPAKGIQKQNLNVEFGDQILSVVIDVAGEEAYHFQRLFGKLIIPDKCRYEVLSTKVEIR 230
QY 236 LAKAEQVTTWTLDSGRPKTVPOKISTPATPRPSPSSKAKKDWKLEAEVKKKEE 295
Db 231 LAKAEIITWASLEYVKGQALLP-KPNVASVQRPVPSKPAKMDKLEAEVKKKEE 289
QY 296 KLDGDAALNKFRRDIYKDAEDMRRAMKSFVBSNGTGLSTNNKDWGAKKVEGSPDGM 355
Db 290 KLDGDAAMNKFSTDIIYQSADEDMRRAMKSFVBSNGTGLSTNNKVEVGTGKVESTPDGM 349
QY 356 LKKWEY 361

```

```

Db          350 LKWEY 355
RESULT 11
Q84LL4      PRELIMINARY; PRT; 273 AA.
ID AC Q84LL4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SGT1-like protein (Fragment).
GN Name=sgt1b;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN R1
RP SEQUENCE FROM N.A.
RX MEDLINE=22882917; PubMed=14504394; DOI=10.1073/pnas.2035132100;
RA Pawlowski J., Holzmann M., Berner C., Fahrni J., Gooday A.J.,
RA Cedhagen T., Habura A., Bowser S.S.;
RT "The evolution of early Foraminifera."
RL Proc. Natl. Acad. Sci. U.S.A. 100:11494-11498 (2003).
RN R2
RP SEQUENCE FROM N.A.
RX Shen W.H.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ318020; CAC85267.1; -.
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR007699; SGS.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
FT NON TER 1
SQ SEQUENCE 273 AA; 30721 MW; 0D8A25034F64DC59 CRC64;

Query Match 46.7%; Score 870; DB 2; Length 273;
Best Local Similarity 59.7%; Pred. No. 9.3e-49;
Matches 169; Conservative 45; Mismatches 55; Indels 14; Gaps 4;

Qy 80 ACIKLEIYQAKALELGSVYASGSRFARLLKCEDEIRAEBSQAPVKNVEATVAATIE 139
Db 2 ACMLEIYRKTALKEGASITPSEKFKLIDECNLFITEEKDL-VQVPSTLPSSV- 59
Qy 140 DKEDFTNMENTPPVIEPPSKP--KYRHDYNSATEVVLTFPAKGVADSVVIDFGEQMLS 197
Db 60 -----TAPPVSELDVTPTAKYRHEYQKPEVVVTVFAKIPKQNVNIDFGEQILS 110
Qy 198 VSIKVPGEHPHFOPRLFSKIIPEKCKYQVLSKVEIRLAKAEQVTTTLDYSGRPKTPV 257
Db 111 VVIEVPGEDAYYLPRLFGKIIPDKCKYEVLSKIEICLAKADITITWASLEHGKGPVLP 170
Qy 258 QKISTPETAAPRPSYPSKAKDWDKLEAEVKKKEEKLGDGAALNKFPRDIYKDADED 317
Db 171 -KPNVSSEQRPAIPSSKKYKWDKLEAEVKKQEKDEKLEGDAALNKFPRDIYQNADED 229
Qy 318 MRRAMKSFVESNGTVLSTNWKVDGAKKVEGSPDGMELKKWE 360
Db 230 MRRAMKSFVESNGTVLSTNWKQEVGTITESTPPDGMELKKWE 272

RESULT 12
SUQT_HUMAN
ID SUQT_HUMAN STANDARD; PRT; 333 AA.
AC Q9Y2Z0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Suppressor of G2 allele of SKP1 homolog (Sgt1) (Putative 40-6-3
DE protein).
GN Name=SGT1;
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN R1
RP SEQUENCE FROM N.A., AND SUBUNIT.
RX MEDLINE=99374050; PubMed=10445024; DOI=10.1016/S1097-2765(00)80184-7;
RA Kitagawa K., Skowrya D., Elledge S.J., Harper J.W., Hieter P.;
RT "SGT1 encodes an essential component of the yeast kinetochore assembly
pathway and a novel subunit of the SCF ubiquitin ligase complex."
RL Mol. Cell 4:21-33(1999).
RN R2
RP SEQUENCE FROM N.A.
RX Schmidt T.;
RL Thesis (2001), University of Goettingen, Germany.
RN R3
RP SEQUENCE FROM N.A.
RX Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RX Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RX Phelan M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator(TM) system donor
vector."
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN R4
RP SEQUENCE FROM N.A.
RX TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Maman A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN R5
RP INTERACTION WITH S100A6.
RX MEDLINE=22741750; PubMed=12746458; DOI=10.1074/jbc.M211518200;
RA Nowotny M., Spiechowicz M., Jastrzebska B., Filipiek A., Kitagawa K.,
RA Kuznicki J.;
RT "Calcium-regulated interaction of Sgt1 with S100A6 (calyculin) and
other S100 proteins."
RL J. Biol. Chem. 278:26923-26928(2003).
CC -!- FUNCTION: May play a role in ubiquitination and subsequent
proteosomal degradation of target proteins.
CC -!- SUBUNIT: Probably associates with SCF (SKP1-CUL1-F-box protein)
complex through interaction with SKP1. Interacts with S100A6.
CC -!- SIMILARITY: Contains 1 CS domain.
CC -!- SIMILARITY: Contains 1 SGS domain.
CC -!- SIMILARITY: Contains 3 TPR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF132856; AAD30062.1; -.
DR EMBL; AJ344097; CAC51433.1; -.
DR EMBL; BT009798; AAP88800.1; -.

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:28:41 ; Search time 134 Seconds
(without alignments)
895.417 Million cell updates/sec

Title: US-10-609-078-8

Perfect score: 1861

Sequence: 1 MAASDLSSKAKEAFVDDDFE.....GAKKVEGPPDGMELKKWEY 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1861	100.0	388	15	US-10-425-114-61162
2	1861	100.0	390	15	US-10-425-114-61204
3	1771.5	95.2	395	16	US-10-767-7701-45515
4	1740	93.5	393	15	US-10-425-114-61121
5	1740	93.5	416	15	US-10-425-114-65836
6	1578	84.8	376	16	US-10-437-963-139712
7	1263	67.9	256	15	US-10-425-114-61200
8	1176.5	63.2	357	15	US-10-424-599-224769
9	1146.5	61.6	359	15	US-10-424-599-208541
10	925.5	49.7	357	15	US-10-424-599-241114
11	729	39.2	160	15	US-10-424-599-261373
12	615	33.0	344	15	US-10-264-049-3093
13	603.5	32.4	223	16	US-10-437-963-146392
					Sequence 61162, A
					Sequence 61204, A
					Sequence 45515, A
					Sequence 65836, A
					Sequence 139712, A
					Sequence 61200, A
					Sequence 224769, A
					Sequence 208541, A
					Sequence 241114, A
					Sequence 261373, A
					Sequence 3093, Ap
					Sequence 146392, A

14	599	32.2	168	15	US-10-424-599-224775	Sequence 224775, A
15	594	31.9	365	14	US-10-287-218-12	Sequence 12, Appl
16	594	31.9	365	16	US-10-474-231-12	Sequence 12, Appl
17	471	25.3	229	16	US-10-437-963-146382	Sequence 146382, A
18	464	24.9	196	10	US-09-983-802-393	Sequence 393, App
19	464	24.9	196	10	US-09-984-490-393	Sequence 393, App
20	464	24.9	196	11	US-09-973-278-503	Sequence 503, App
21	464	24.9	228	10	US-09-983-802-401	Sequence 401, App
22	464	24.9	228	10	US-09-984-490-401	Sequence 401, App
23	464	24.9	228	11	US-09-973-278-511	Sequence 511, App
24	445	23.9	413	14	US-10-032-585-7066	Sequence 7066, Ap
25	442.5	23.8	123	15	US-10-424-599-241112	Sequence 241112, A
26	405	21.8	400	14	US-10-128-714-3417	Sequence 3417, Ap
27	403	21.7	123	15	US-10-424-599-218698	Sequence 218698, A
28	383	20.6	478	14	US-10-128-714-8417	Sequence 8417, Ap
29	239	12.8	82	15	US-10-424-599-192478	Sequence 192478, A
30	235.5	12.7	483	16	US-10-437-963-195520	Sequence 195520, A
31	226.5	12.2	482	15	US-10-424-599-191442	Sequence 191442, A
32	224	12.0	482	14	US-10-298-638-27	Sequence 27, Appl
33	224	12.0	485	14	US-10-298-638-10	Sequence 10, Appl
34	223	12.0	533	14	US-10-298-638-28	Sequence 28, Appl
35	219.5	11.8	163	16	US-10-767-701-41397	Sequence 41397, A
36	219.5	11.8	488	15	US-10-425-114-61206	Sequence 61206, A
37	213	11.4	101	15	US-10-424-599-183984	Sequence 183984, A
38	203.5	10.9	955	9	US-09-991-496-127	Sequence 127, App
39	197.5	10.6	1427	9	US-09-874-923-97	Sequence 97, Appl
40	197.5	10.6	1427	9	US-09-991-496-97	Sequence 97, Appl
41	197.5	10.6	1641	9	US-09-874-923-96	Sequence 96, Appl
42	197.5	10.6	1641	9	US-09-991-496-96	Sequence 96, Appl
43	192	10.3	494	14	US-10-043-487-296	Sequence 296, App
44	190	10.2	590	15	US-10-425-114-55903	Sequence 55903, A
45	186	10.0	191	16	US-10-767-701-35628	Sequence 35628, A

ALIGNMENTS

RESULT 1

US-10-425-114-61162
; Sequence 61162, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61162
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-061-D10_FLI.pap
US-10-425-114-61162

Query Match 100.0%; Score 1861; DB 15; Length 388;
Best Local Similarity 100.0%; Pred. No. 3.1e-137;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAASDLSSKAKEAFVDDDFELATLYSQAIIDAGPATADLYADRAQAHKLGNYTEAVADA	60
Db	28	MAASDLSSKAKEAFVDDDFELATLYSQAIIDAGPATADLYADRAQAHKLGNYTEAVADA	87
Qy	61	NKALELDPMMHKAYYRKGAAACIKLEEVQTKAAALELGSYGSGSRFARLLKCDERTAE	120
Db	88	NKALELDPMMHKAYYRKGAAACIKLEEVQTKAAALELGSYGSGSRFARLLKCDERTAE	147

Qy	121	ESSQAPVKVNEATVAATTIEDKEDFTNNMENTPPIVIEPPSPKPYRHDYNSATEVVLTFPAK	180
Db	148	ESSQAPVKVNEATVAATTIEDKEDFTNNMENTPPIVIEPPSPKPYRHDYNSATEVVLTFPAK	207
Qy	181	GYPADSVVIDFGEQMLSVSIEVPGEEPVHFOPRIFSIIPEKCKYQVLSIKVEIRLAKAE	240
Db	208	GYPADSVVIDFGEQMLSVSIEVPGEEPVHFOPRIFSIIPEKCKYQVLSIKVEIRLAKAE	267
Qy	241	QVTTWTLDYSGRPKTVPOKISTPAPTAPRPPSPSSKAKDWDKLEAEVKKEEKEEKLDGD	300
Db	268	QVTTWTLDYSGRPKTVPOKISTPAPTAPRPPSPSSKAKDWDKLEAEVKKEEKEEKLDGD	327
Qy	301	AALNKEFFRDIYKDAEDMRRAMMKSFVESNGTVLSTTNWKDVGAKKVEGSPDGMELKKWE	360
Db	328	AALNKEFFRDIYKDAEDMRRAMMKSFVESNGTVLSTTNWKDVGAKKVEGSPDGMELKKWE	387
Qy	361	Y 361	
Db	388	Y 388	

RESULT 2

```

US-10-425-114-61204
; Sequence 61204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61204
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3732-024-C2_FLI.pep
US-10-425-114-61204

```

```

330 AALNKFRDIYKDADEMRAMMKSFYVESNGTVLSTWVKDVGAKKVBSPPDGMELKKWE 389

Qy      361 Y 361
      |
Db      390 Y 390

RESULT 3
US-10-767-701-45515
; Sequence 45515, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45515
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(395)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2718_1.pep
US-10-767-701-45515

```

RESULT 4

RESULTS 4
US-10-425-114-61212
; Sequence 61212, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 61212
LENGTH: 393
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3960-013-B4_FLI.pap
US-10-425-114-61212

Query Match 93.5%; Score 1740; DB 15; Length 393;
Best Local Similarity 93.1%; Pred. No. 9.3e-128;
Matches 335; Conservative 10; Mismatches 15; Indels 0; Gaps 0;
Qy 1 MAASDLESKAKEAFVDDDFELATLYSOAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
Db 33 MAASDLESKAKEAFVDDDFELAAELYTQAI DAGPATADLYADRAQAHIKLGNYTEAVADA 92
Qy 61 NKAIELDPMMHKAYYRKGAAACIKLEEYQTAKAALGSSYASGDSRFRLLKCEDECIAE 120
Db 93 NKAIGLDPTHMKAYYRKGAAACIKLEEYQTAKAALGSSYAPGDSRFRLLKCEDECIAE 152
Qy 121 ESSQAPKVNVEATVAATIEDKEDFTNMTPPVIEPPSKPKYRHDYNSATEVVLTIYAK 180
Db 153 ESSQAPAKNVEAPVAATVEDKEDVANMDNTPPVVEPPSKPKYRHDYNSATEVVLTIYAK 212
Qy 181 GVPADSVVIDFGQMLSVSIEVPGEYPHFQPRFLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
Db 213 GVPADSVVIDFGQMLSVSIEVPGEYPHFQPRFLFSKIIPEKCKYQVLSTKVEIRLAKAE 272
Qy 241 QVTWTLIDYSGRPKTPVKISTPAETAPRPSYPSKAKDWDKLEAEVKKKEEKLGD 300
Db 273 QVTWTLIDYSGRPKTPVKISTPAETAPRPSYPSKAKDWDKLEAEVKKKEEKLGD 332
Qy 301 AALNKFFRDIYKDAEDMRRAMKSFVSNGTVLSTNWKDYGAKKVGSPDGMELKKWE 360
Db 333 AALNKFFRDIYKDAEDMRRAMKSFRESNGTVLSTNWKDVGSKTVEASPPDGMELKKWE 392

RESULT 5
US-10-425-114-65836
Sequence 65836, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 65836
LENGTH: 416
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3591-005-E5_FLI.pap
US-10-425-114-65836

Query Match 93.5%; Score 1740; DB 15; Length 416;

Best Local Similarity 93.1%; Pred. No. 1e-127;
Matches 335; Conservative 10; Mismatches 15; Indels 0; Gaps 0;
Qy 1 MAASDLESKAKEAFVDDDFELATLYSOAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
Db 56 MAASDLESKAKEAFVDDDFELAAELYTQAI DAGPATADLYADRAQAHIKLGNYTEAVADA 115
Qy 61 NKAIELDPMMHKAYYRKGAAACIKLEEYQTAKAALGSSYASGDSRFRLLKCEDECIAE 120
Db 116 NKAIGLDPTHMKAYYRKGAAACIKLEEYQTAKAALGSSYAPGDSRFRLLKCEDECIAE 175
Qy 121 ESSQAPKVNVEATVAATIEDKEDFTNMTPPVIEPPSKPKYRHDYNSATEVVLTIYAK 180
Db 176 ESSQAPAKNVEAPVAATVEDKEDVANMDNTPPVVEPPSKPKYRHDYNSATEVVLTIYAK 235
Qy 181 GVPADSVVIDFGQMLSVSIEVPGEYPHFQPRFLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
Db 236 GVPADSVVIDFGQMLSVSIEVPGEYPHFQPRFLFSKIIPEKCKYQVLSTKVEIRLAKAE 295
Qy 241 QVTWTLIDYSGRPKTPVKISTPAETAPRPSYPSKAKDWDKLEAEVKKKEEKLGD 300
Db 296 QVTWTLIDYSGRPKTPVKISTPAETAPRPSYPSKAKDWDKLEAEVKKKEEKLGD 355
Qy 301 AALNKFFRDIYKDAEDMRRAMKSFVSNGTVLSTNWKDYGAKKVGSPDGMELKKWE 360
Db 356 AALNKFFRDIYKDAEDMRRAMKSFRESNGTVLSTNWKDVGSKTVEASPPDGMELKKWE 415

RESULT 6
US-10-437-963-139712
Sequence 139712, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 139712
LENGTH: 376
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_40977C.1.pap
US-10-437-963-139712

Query Match 84.8%; Score 1578; DB 16; Length 376;
Best Local Similarity 82.1%; Pred. No. 4.2e-115;
Matches 307; Conservative 27; Mismatches 24; Indels 16; Gaps 3;
Qy 2 AASDLESKAKEAFVDDDFELATLYSOAIDAGPATADLYADRAQAHIKLGNYTEAVADA 61
Db 5 AASDLESKAKEAFVDDDFELAAELYTQAI EASPATADLYADRAQAHIKLGNYTEAVADA 64
Qy 62 KAIELDPMMHKAYYRKGAAACIKLEEYQTAKAALGSSYASGDSRFRLLKCEDECIAE 121
Db 65 KAIELDPMMHKAYYRKGAAACIKLEEYQTAKAALGSSYASGDSRFRLLKCEDECIAE 124
Qy 122 SSQAPKVNVE-----ATVAATIEDKEDFTNMTPPVIEPPSKPKYRHDYNSATEVVLTI 176
Db 125 LSEVPVKKAEDGAAAPSVASVFEKDDAANMDNTPPMVE--VKPKYRHDYNSATEVVLTI 182
Qy 177 IFAGKVPADSVVIDFGQMLSVSIEVPGEYPHFQPRFLFSKIIPEKCKYQVLSTKVEIRL 236

Db 183 IFAGVPAENVVDFGEQMLSVSIEVPGPEYHFPQRLFSKIIPEKSYQVLSTKVEIRL 242
Qy 237 AKAEQVTTTLDYSGRPKTVPOKISTP-----AETAPRPSYPSSKAKDWDKLEAE 287
Db 243 AKAEQVTTTLDYSGRPKTVPOKISTP-----AETAPRPSYPSSKAKDWDKLEAE 302
Qy 288 VKKEEKEEKLGDGAALNKFRRDIYKDAEDMRRAMKSFVESNGTVLSTNNKWDVGAKVE 347
Db 303 VKKEEKEEKLGDGAALNKFRRDIYKDAEDMRRAMKSFVESNGTVLSTNNKWDVGSKVE 362
Qy 348 GSPDPGMELKKWEY 361
Db 363 GSPDPGMELKKWEY 376

RESULT 7

US-10-425-114-61200
; Sequence 61200, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(531313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 61200
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-210-C7_FLI.pap
US-10-425-114-61200

Query Match 67.9%; Score 1263; DB 15; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.1e-90;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 120 BSSQAPVKVNEATVAATIEDKEDFTNMENPPVIEPPSKYRHYDYNSTATEVLTIFA 179
Db 15 BSSQAPVKVNEATVAATIEDKEDFTNMENPPVIEPPSKYRHYDYNSTATEVLTIFA 74
Qy 180 KGVPADSVVIDFGQMLSVSIEVPGPEYHFPQRLFSKIIPEKCKYQVLSTKVEIRLAKA 239
Db 75 KGVPADSVVIDFGQMLSVSIEVPGPEYHFPQRLFSKIIPEKCKYQVLSTKVEIRLAKA 134
Qy 240 EQVTTTLDYSGRPKTVPOKISTP-----AETAPRPSYPSSKAKDWDKLEAEVKEEKEEKLGD 299
Db 135 EQVTTTLDYSGRPKTVPOKISTP-----AETAPRPSYPSSKAKDWDKLEAEVKEEKEEKLGD 194
Qy 300 DAALNKFRRDIYKDAEDMRRAMKSFVESNGTVLSTNNKWDVGAKVEGSPDGMELKKW 359
Db 195 DAALNKFRRDIYKDAEDMRRAMKSFVESNGTVLSTNNKWDVGAKVEGSPDGMELKKW 254
Qy 360 EY 361
Db 255 EY 256

RESULT 8

US-10-424-599-224769
; Sequence 224769, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 224769
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(357)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_44998C.1.pap
US-10-424-599-224769

Query Match 63.2%; Score 1176.5; DB 15; Length 357;
Best Local Similarity 64.5%; Pred. No. 1e-83;
Matches 234; Conservative 48; Mismatches 68; Indels 13; Gaps 6;
Qy 3 ASDLESKAKAFVDDDELATYLSQAIDAGPATADLYADRAQAHIKIGNYTEAVADANK 62
Db 2 ASDLESKAKAFVDDDELATYLSQAIDAGPATADLYADRAQAHIKIGNYTEAVADANK 61
Qy 63 AIELDPMHMKAYKGAACIKLEBYOTAKALELGSYASGDSFARLLKCEDERIAEES 122
Db 62 AIELNPSHSAKYLKRGKACIKLEBYOTAKALELGSYASGDSFARLLKCEDERIAEES 121
Qy 123 SQAPVKVNEATVAATIEDKEDFTNMENPPVIEPPS-----KPKYRHYDYNSTATEVLTIF 178
Db 122 GVPIQESTTQGAATKAVE-----AENDLP--EPPTVVVKYRHFYQKPDENVITIF 175
Qy 179 AKGVPADSVVIDFGQMLSVSIEVPGPEYHFPQRLFSKIIPEKCKYQVLSTKVEIRLAK 238
Db 176 AKGIPRDSITVDFGEQILSVTINIPCKDAYVFPQRLFGKIIPSKRYEVLSTKIEICLAK 235
Qy 239 AEQVTTTLDYSGRPKTVPOKISTP-----AETAPRPSYPSSKAKK--DWDKLEAEVKEEKEEK 296
Db 236 ADHIQWTSLEFN--KGSTVAQRFSVLVARGEKPTYPSSKPKKDDWDKLEAQVKEEKEEK 294
Qy 297 LDGDAALNKFRRDIYKDAEDMRRAMKSFVESNGTVLSTNNKWDVGAKVEGSPDGMEL 356
Db 295 LDGDAALNKFRRDIYKDAEDMRRAMKSFVESNGTVLSTNNKWDVGAKVEGSPDGMEL 354
Qy 357 KKW 359
Db 355 KKW 357

RESULT 9

US-10-424-599-208541
; Sequence 208541, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 208541
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_30341C.1.pap

US-10-424-599-208541

Query Match 61.6%; Score 1146.5; DB 15; Length 359;
Best Local Similarity 62.8%; Pred. No. 2.3e-81;
Matches 230; Conservative 55; Mismatches 66; Indels 15; Gaps 5;

QY 3 ASDLESKAEAFVDDDFELATLEYLSQAIDAGPATADLYADRAQAHIKLGNTYEAVADANK 62
DB 2 ASDLEAKAEAFVDDDFELAVDLISQAHLFENKAEFYADRAQANIKLNNTYEAVADANK 61
QY 63 AIELDPMMHKAYRKGACIKLEEVQYAKALELSSYASGDSRPARLLKCEDERIAEES 122
DB 62 AIELNPSLPKAYLRKGTAKMKEVETAKALEVGASLSPNSRFATLIKCDKLIABES 121
QY 123 SQAPKVNVEATVAATIEDKEDFTNMNENTPPVIEPPSKYRHDYNSATEVVLTK 176
DB 122 YTPPIIEKATTT-----EDVTQKVOQDDLEKPTVTVTVPKRYHEFYQKHDQVVVT 174
QY 177 IFAGKVPADSVVDFGEQMLSVSIEVPGEPYHFQPLRFKIIPEKCKYQVLSTKVEIRL 236
DB 175 IFAPKIPKESITVDFGEQILSVSINVGTEDYVQPLRFGLIIFNSNCYEVLSKIEIRL 234
QY 237 AKAEQVTTWTLTDSGRPKTPVKQISTPAET-APRPSYSSKAKDMDKLEAEVKEEKEE 295
DB 235 AKABPIHTWLEFT-RDSVWQQRVSALSVTGGKPTYPSSKRCENTTEGEAQVKKEGDE 293
QY 296 KLDGDAALNKFRDIYKDADEDMRRAMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGME 355
DB 294 KLDGDAALNKFRDIYQDADEDTTRAMKSFVESNGTVLSTNWKVEVGTKKVEGSPPDGME 353
QY 356 LKKWEY 361
DB 354 LKKWEY 359

RESULT 10

US-10-424-599-241114
; Sequence 241114, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 241114
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)..(357)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59754C.1.pep
US-10-424-599-241114

Query Match 49.7%; Score 925.5; DB 15; Length 357;
Best Local Similarity 53.3%; Pred. No. 4.5e-64;
Matches 194; Conservative 53; Mismatches 104; Indels 13; Gaps 5;

QY 3 ASDLESKAEAFVDDDFELATLEYLSQAIDAGPATADLYADRAQAHIKLGNTYEAVADANK 62
DB 2 ATALEKKAKEAFFDDDFGLAVLDYSEAIRLDFNDANLFAADRAQAHIKLNFAEVSADANK 61
QY 63 AIELDPMMHKAYRKGACIKLEEVQYAKALELSSYASGDSRPARLLKCEDERIAEES 122
DB 62 SIQLNPSLPKAYLRKATACIKLEQYHTAKVALQNGAFAQDSDRPFANLIQQDCRCIAEYE 121

QY 123 SQAPKVNVEATVAATIEDKEDFTNMNENTPPVIEPPSKYRHDYNSATEVVLTKFAKV 182
DB 122 SS-----ERLLAPCLPHLSNKKSWDEKKDKETVCPKRMPEHYTNPCEGSNTTXHRNL 175
QY 183 PADSWIDFGEQMLSVS---IEVPGEPYHFQPLRFS-KIIPKCKYQVLSTKVEIRLAK 238
DB 176 DQDGAFFMFKLQLQMLHQLQVKDEHSQHVFPCLFGYMIIPNNCRVEVLSKIEIRLAK 235
QY 239 AEQVTTWTLTDSGRPKTPVKQISTPAETAPRPSYSSKAK-KOWDKLEAEVKEEKEEKL 297
DB 236 AEANWTSLEYG--KNTLPPIINRPVQSERASYSPKPRTKDWDKLEAQVKCEKEEKL 293
QY 298 DGDAAALNKFRDIYKDADEDMRRAMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMEKL 357
DB 294 DGDAAALNKFRDIYQNADEDMRRAMKSFLESNGTVLSTDMKVEVGSKKVEGSPPEGMEKL 353
QY 358 KWEY 361
DB 354 KWEY 357

RESULT 11

US-10-424-599-261373
; Sequence 261373, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 261373
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)..(160)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78041C.1.pep
US-10-424-599-261373

Query Match 39.2%; Score 729; DB 15; Length 160;
Best Local Similarity 89.2%; Pred. No. 3.6e-49;
Matches 140; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 97 GSSVASGDSRPARLLKCEDERIAEESQAPKVNVEATVAATIEDKEDFTNMNENTPPVIEP 156
DB 4 GSSYAPGDSRPTRLKCEDECIAEESQAPKVNVEATVAATIEDKEDVANMDNTPPVVEP 63
QY 157 PSKPKYRHDYNSATEVVLTKFAKVGPADSVVDFGEQMLSVSIEVPGEPYHFQPLRFS 216
DB 64 PSKPKYTHDYNSATEVVLTKYAKGVPADSVVDCGQMLSVSIEVPGEPYHFQXPRELS 123
QY 217 KIIPKCKYQVLSTKVEIRLAKAEQVTTWTLTDSGRP 253
DB 124 KIIPKCKYHLSTKVEIRLAKAEQVTTWTLTDSGRP 160

RESULT 12

US-10-264-049-3093
; Sequence 3093, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3093
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3093

Query Match          33.0%; Score 615; DB 15; Length 344;
Best Local Similarity 38.7%; Pred. No. 8.6e-40;
Matches 135; Conservative 62; Mismatches 116; Indels 36; Gaps 5;

QY 12 EAFVDDDELATYLSQADPATADYADRAQAHIKLGNYTEAVADANKAIELDPMH 71
DB 31 DALIDEDPQAALBELTKALEQPDPAQYQCQAYCHILLGNYCVAVADAKSLNPNNS 90

QY 72 KAYVRKGAACIKLEBYQYAKAALLEGSSYAGDSRPFARLLKCEDERIAESSQAPVKNVE 131
DB 91 TAMLKRGICEYHEKNYAAALETFTEGQKLDSDANFVMIKRCQE--AQNGSESEV--- 144

QY 132 ATVAATIEDKEDFTNMENTPVPVPPSPKPYRHDYNSATEVLTIFAKGVPADSVVIDF 191
DB 145 -----WTH-----QSKIKDYMYQTESQVWITLMKNQKNDVNVF 180

QY 192 GEQMLSVIEVPGPEYHFQRLFSKIIPKCKYQVLSTKVEIRLAKAEQVTTTLTLDYSG 251
DB 181 SEKELSALVKLPSGEDYNLKLELHPILPEOSTFKVLSTKIEIKLKEPAVRWEKLEQ 240

QY 252 RPKVTPKISTPAETAPRPSPSKA-KKWDKLEAEVKEKEKLEKLDGDAALNKFREDI 310
DB 241 DVPTPKQFVAD-----VKNLYPSSSPYTRNWDKLVGEIKKEKNEKLEKLDGDAALNRLFO 295

QY 311 YKDAEDMRRAMKSFVESNGTIVLSTNNKVDGAKKVEGSPDGMELKKW 359
DB 296 YSGSDEVKRAMKSFVESGGTIVLSTNNKVDGAKKVEINPPDDMEWKY 344

RESULT 13
; Sequence 146392, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146392
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47020C.1.pep
US-10-437-963-146392

Query Match          32.4%; Score 603.5; DB 16; Length 223;
Best Local Similarity 47.5%; Pred. No. 3.8e-39;
```

```
Matches 125; Conservative 45; Mismatches 44; Indels 49; Gaps 7;

QY 101 ASGDSREARLLKCEDERIAESSQAPVKNVEATVAATIEDKEDFTNMENTPVPVIEPSKP 160
DB 8 APGDARFAKLAKELDDKAPKAAAAPPAASEDQAAAS-----SGKP 48

QY 161 KYRHDYNSATEVLTIFAKGVPADSVVIDFGEQMLSVIEVPGPEYHFQRLFSKIIP 220
DB 49 KYKDFNGASEVVTTFVFAKVAEHVSVDVGEQMLSVSEV-----VP 92

QY 221 EKCKYQVLSTKVEIRLAKA-EQVTTTLTLDYSGRPTVPQKISTPAETAPRPSYSSKAKK 279
DB 93 AKCRYSILPSKIEVRLAKADEQVTTSLVTSKAN---NKLAATATT-----TRKKV 142

QY 280 DWDKLEAEVKEKEKLEKLD-GDAALNKFRRDIYKDADEDMRRAMKSFVESNGTIVLSTNW 338
DB 143 DWDKLEAEVKEKEKEEVEVDATPVVNRFFQOMYNGDEDMRRAIMKSYVES--YVLSTDW 200

QY 339 KDVGAKKVEGSPDGMELKKWEY 361
DB 201 KDVGSKKIEASAPGEMELKKWEY 223

RESULT 14
US-10-424-599-224775
; Sequence 224775, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224775
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(168)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45001C.1.pep
US-10-424-599-224775

Query Match          32.2%; Score 599; DB 15; Length 168;
Best Local Similarity 65.9%; Pred. No. 5.7e-39;
Matches 116; Conservative 20; Mismatches 22; Indels 18; Gaps 2;

QY 196 LSVSIEVPGPEYHFQRLFSKIIPKCKYQVLSTKVEIRLAKAEQVTTTLTLDYSGRPT 255
DB 1 LSVTINIPCKDAVVFQRLFGKIIPSKRCYEVLTSTKIEICLAKTDHQTSLDFN----- 55

QY 256 VPQKIST-----PAPTAPRPSYSSKAKKDWKLEAEVKEKEKLEKLDGDAALNK 305
DB 56 ---KASTDILHRLGMLLQLLEVXRPSYSSKQTRDWDKIEAQVKKEKLEKLDGDAALNK 112

QY 306 FFRDIYKDADEDMRRAMKSFVESNGTIVLSTNNKVDGAKKVEGSPDGMELKKWEY 361
DB 113 FFREIQDADEDMRRAMKSFVESNGTIVLSTNNKVEGSKKVEGSPDGMELKKWEY 168

RESULT 15
US-10-287-218-12
; Sequence 12, Application US/10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
```


This Page Blank (uspto)

=> file biosis caplus caba agricola
=> s rar1
L1 165 RAR1

=> duplicate remove 11
L2 80 DUPLICATE REMOVE L1 (85 DUPLICATES REMOVED)

=> d t1 1-25

L2 ANSWER 1 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 Calcium-independent phospholipase A2 is regulated by a novel protein kinase C in human coronary artery endothelial cells
L2 ANSWER 2 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on STN
T1 Mammalian CHORD-containing protein 1 is a novel heat shock protein 90-interacting protein.

L2 ANSWER 3 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 RAR1, ROR1, and the actin cytoskeleton contribute to basal resistance to Magnaporthe grisea in barley

L2 ANSWER 4 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 Barley Rom1 reveals a potential link between race-specific and nonhost resistance responses to powdery mildew fungi

L2 ANSWER 5 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 The atypical resistance gene, RPM8, recruits components of basal defence for powdery mildew resistance in Arabidopsis

L2 ANSWER 6 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 Pesticidal cyanine dye derivatives

L2 ANSWER 7 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on STN
T1 RAR1 positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance.

L2 ANSWER 8 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on STN
T1 Arabidopsis RIN4 negatively regulates disease resistance mediated by RPS2 and RPM1 downstream or independent of the NDR1 signal modulator and is not required for the virulence functions of bacterial type III effectors AvrPpt2 or AvrPpm1.

L2 ANSWER 9 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 Cutting edge transcriptome analysis: It's all about design
L2 ANSWER 10 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on
T1 Molecular chaperone Hsp90 associates with resistance protein N and its signaling proteins SGT1 and Rar1 to modulate an innate immune response in plants.

L2 ANSWER 11 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 Prothrombin overexpressed in postnatal neurons requires blood factors for activation in the mouse brain

L2 ANSWER 12 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on
T1 Signaling requirements and role of salicylic acid in HRT- and rrt-mediated resistance to turnip crinkle virus in Arabidopsis.

L2 ANSWER 13 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
T1 Mammalian CHORD-containing protein 1 is a novel heat shock protein 90-interacting protein

L2 ANSWER 14 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on

This Page Blank (uspto)